Structural basis for signaling in innate immunity via autocatalytical changes of MAVS by NMR

Von der Fakultät für Lebenswissenschaften

der Technischen Universität Carolo-Wilhelmina

zu Braunschweig

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

genehmigte

D i s s e r t a t i o n

von Lichun He
aus Shanxi, China

1. Referentin: Prof.Dr.Christiane Ritter
2. Referent: Prof.Dr. Stefan Dübel
eingereicht am: 11.06.2014
mündliche Prüfung (Disputation) am: 22.09.2014

Druckjahr 2015

Vorveröffentlichungen der Dissertation

Teilergebnisse aus dieser Arbeit wurden mit Genehmigung der Fakultät für Lebenswissenschaften, vertreten durch die Mentorin der Arbeit, in folgenden Beiträgen vorab veröffentlicht:

Tagungsbeiträge

He, L. & Ritter, C.: Characterization of the structural basis for altered host protein trafficking by the adenoviral protein E4orf3. (Poster) 5th International PhD Symposium, Braunschweig (2011)

He, L. & Ritter, C.: Characterization of the structural basis for altered host protein trafficking by the adenoviral protein E4orf3. (Poster) EMBO Practical Course in Joachimsthal, Berlin (2012)

He, L. Ahmed, M. & Ritter, C.: Structural basis for signaling in innate immunity via autocatalytic changes of MAVS by solid state NMR. (Poster) European Network on Spin Hyperpolarization Physics and Methodology in NMR and MRI in Hersonissos, Crete, Greece (2013)

He, L. Ahmed, M. Spehr, J & Ritter, C.: Structural basis for signaling in innate immunity via autocatalytic changes of MAVS by solid state NMR. (Poster) 6th International PhD Symposium, Braunschweig (2013)

Acknowledgements

My doctoral would not have been accomplished without the help from several people. I would like to dedicate my sincere thanks to them.

First, I would like to thank my advisor, Prof. Dr. Christiane Ritter, who lead me entering the NMR world. She has given me scientific guidance, support, encouragement; overlooked many of my mistakes and helped me to be a confident Ph.D. student. Her patient instruction, friendly company in all the three and half years helped me to complete the thesis work. I have improved myself a lot in her research group, being one of her dedicated team and working in a creative environment. I would not have been possible to write the thesis without her guidance.

I would also like to thank Dr. Mumdooh Ahmed for his tremendous help with recording all the solid state NMR spectra and teaching on NMR data processing and analysis. He also explained complicated NMR theories in a simple way and helped me to understand them. His kind assistance, productive discussion and brainstorming about the project helped me to complete the projects.

I am also very grateful to Dr. Johannes Spehr and Dr. Tobias Schubeis. Thanks to their helps on teaching me about data acquirement, processing and analysis in the beginning of my work in the lab. Dr. Tobias Schubeis taught me how to use CcpNmr Analysis program. Dr. Johannes Spehr helped a lot in NOESY spectra analysis and structural calculation by UNIO. In addition, Their help on translation of German documents made my life much easier in Germany.

I sincerely thank all my thesis committee members Dr. Joop van den Heuvel and Prof. Dr. Melanie M. Brinkmann for providing scientific discussions and advices.

It is my great pleasure to thank my collaborators Dr. Benjamin Bardiaux, Dr. Heinrich Lünsdorf, Dr. Renate König, Prof. Dr. Louise Serpell and Dr. Ulfert Rand. It was a perfect collaboration and I have learned a lot from all of them. I would like to point out that Dr. Benjamin Bardiaux contributed remarkably on structural calculation by ARIA, including providing me new version of ARIA, teaching me how to use it and doing a lot of structural calculation work. Dr. Heinrich Lünsdorf recorded all the EM images for me and helped me on analysis of the EM images.

My sincere thanks also go to Puwei Yuan and Jan Hellert. Thanks for their constructive discussions and suggestions. Thanks to all the lab members for the friendly environment in the lab. I would also like to thank all my friends in Braunschweig, for helping in my daily life and making me not feel alone in abroad.

Acknowledgements

Last, but not least, my thanks and love go to my wife Lisha Zha, for her persisting help, her loving support and feeling happy or sad together with me.

I sincerely thank all of you.

Lichun He

Braunschweig, 2014

Contents

1 INTRODUCTION 1

1.1 PATHOGEN RECOGNITION AND INNATE IMMUNITY 1

1.1.1 Pathogen product recognition by innate immune system 1

1.1.2 Toll like receptors involved in the detection of nucleic acids 2

1.1.3 RIG-I like receptors and their regulation in antiviral immunity 2

1.1.3.1 Members of RLR family 3

1.1.3.2 Distinct recognition pattern of RLRs 4

1.1.3.3 Signal pathway of RLR mediated immune response 5

1.1.3.4 Regulation of RLR signaling cascade 6

1.1.4 MAVS – an essential adaptor in RLR antiviral pathway 7

1.1.4.1 MAVS in type I interferon production and inflammasome activation 8

1.1.4.2 Prion-like behaviour of MAVS in RIG-I signaling 9

1.1.4.3 Regulation of MAVS by mitochondrial membrane and proteins 10

1.1.4.3.1 Regulation through protein-protein interaction 10

1.1.4.3.2 Regulation through mitochondria dynamic and membrane potential 12

1.1.4.3.3 Regulation through ubiquitination and phosphorylation of MAVS 13

1.1.4.4 Viral antagonism of MAVS mediated signalling 14

1.1.5 Structural studies on CARD and other DD members 15

1.1.6 High-order assemblies as platform for signal transduction 17

1.2 NUCLEAR MAGNETIC RESONANCE (NMR) SPECTROSCOPY 17

1.2.1 Elementary principles of NMR spectroscopy 18

1.2.2 Nuclear magnetic spins interaction 20

1.2.2.1 Chemical shift 20

1.2.2.2 Scalar coupling 22

1.2.2.3 Dipole-dipole interaction 22

1.2.3 Solution NMR spectroscopy and structure determination 23

1.2.3.1 Sample preparation and spectra measurements 23

1.2.3.2 Assignments and structure calculation 24

1.2.4 Relaxation and dynamics in solution 24

1.2.4.1 Relaxation 25
1.2.4.2 Chemical exchange 25
1.2.5 Solid-state NMR spectroscopy 26
1.2.5.1 Magical angle spinning 26
1.2.5.2 Cross polarization 27
1.2.5.3 Recoupling in MAS solid state NMR 29
1.2.6 Protein structure determination by solid state NMR 29
1.2.6.1 Labeling strategy 30
1.2.6.2 3D structure determination by solid state NMR 31
1.3 CONTRIBUTION FROM OTHER TECHNIQUES TO PROTEIN STRUCTURE DETERMINATION 32
1.3.1 Electron microscopy 32
1.3.1.1 Structural information available from electron microscopy 33
1.3.1.2 3D reconstruction of helical assemblies 33
1.4 AIM OF THE PROJECT 36

2 MATERIALS AND METHODS 37

2.1 MATERIAL 37
2.1.1 Enzymes, Kits and Molecular weight standards 37
2.1.2 Plasmids 38
2.1.3 Oligonucleotides 40
2.1.4 Bacterial strains 41
2.1.5 Media and Buffer 42
2.2 METHODS 43
2.2.1 Cloning and mutagenesis 44
2.2.2 Protein production and purification 44
2.2.3 Protein analyse and characterization 46
2.2.3.1 SDS polyacrylamide gel electrophoresis 46
2.2.3.2 Mass spectrometry and circular dichroism 46
2.2.3.3 Pull down assay 47
2.2.4 Solution state NMR 47
2.2.4.1 Data acquisition and processing 47
2.2.4.2 Backbone and side chain resonance assignment 48

2.2.4.3 Chemical shift index for secondary structure 49
2.2.4.4 Structural calculation 50
2.2.5 Solid state NMR 50
2.2.5.1 Data acquisition and processing 50
2.2.5.2 Resonance assignment 52
2.2.5.3 Structure determination by solid state NMR 52
2.2.6 Negative staining electron microscopy 53
2.2.6.1 Sample preparation 53
2.2.6.2 Energy-filtered transmission electron microscopy 53
2.2.7 Electron microscopy images processing 53
2.2.8 In vivo cell experiment 54
2.2.8.1 Electroporation 54

3 RESULTS 56

3.1 MONOMERIC WT MAVS-CARD STRUCTURE AND DYNAMICS BY SOLUTION NMR 56
3.1.1 Expression and purification of MAVS and RIG-I 56
3.1.2 Morphology of MAVS-CARD$_{1-100}$ in different buffers 58
3.1.3 Secondary structure and intrinsic fluorescence of MAVS-CARD$_{1-100}$ 59
3.1.4 Chemical shift assignment of MAVS-CARD$_{1-100}$ at pH 3.0 60
3.1.5 Structure calculation of monomeric MAVS-CARD$_{1-100}$ by solution NMR 63
3.1.6 Surface characteristics 64
3.2 MAVS-CARD$_{1-100}$ MUTANTS AND THEIR INTERACTIONS WITH RIG-I CARD 66
3.2.1 Mutants of MAVS-CARD$_{1-100}$ 66
3.2.2 Backbone resonance assignment of the monomeric mutant D23S E26S at pH 7.0 67
3.2.3 Dynamics of MAVS-CARD$_{1-100}$ 69
3.2.4 Reduced spectral density mapping 70
3.2.5 Model free analysis 71
3.2.6 Interaction of MAVS-CARD$_{1-100}$ mutants with RIG I CARD domain 73
3.3 FIBRILLAR MAVS-CARD$_{1-100}$ STRUCTURE BY SOLID STATE NMR AND EM 76
3.3.1 Seeding ability of the in-vitro expressed MAVS-CARD$_{1-100}$ 76
3.3.2 Resonance assignments of the MAVS-CARD$_{1-100}$ filament by solid state NMR 77
3.3.2.1 Spectra quality 78

3.3.2.2 Sequence specific resonance assignment 79
3.3.2.3 Secondary structure from the chemical shift index (CSI) 83
3.3.2.4 Chemical shift perturbation between monomer and filament state MAVS-CARD 84
3.3.3 Structural calculation of MAVS-CARD$_{1-100}$ within the filament 85
3.3.3.1 Distance constraints from uniformly labeled sample 86
3.3.3.2 Distance constraints from sparsely labeled sample 89
3.3.4 De novo structure calculation of MAVS-CARD$_{1-100}$ in fibrillar state by ARIA 92
3.4 ASSEMBLY OF MAVS-CARD$_{1-100}$ FILAMENT 96
3.4.1 Intermolecular distance constraints and three types of interaction 96
3.4.2 Probing surface accessibility of MAVS-CARD filament using paramagnetic relaxation 98
3.4.3 Mutagenesis of the residues on the interaction surface 101
3.5 QUATERNARY STRUCTURE OF MAVS-CARD$_{1-100}$ FILAMENT 102
3.5.1 Symmetry information from electron microscopy images 103
3.5.2 Quaternary structure of MAVS-CARD$_{1-100}$ filament calculated by ARIA 104
4 DISCUSSION AND PERSPECTIVES 107
4.1 MAVS LOCATION ON THE OUTER MEMBRANE OF THE MITOCHONDRIA 107
4.2 RIG-I CARD MIGHT SERVE AS A NUCLEATOR FOR MAVS SELF-ASSEMBLY 107
4.3 CHARACTERIZATION OF MAVS-CARD FILAMENT 109
4.4 THE OLIGOMERIZATION OF MAVS CARD IS IMPORTANT FOR EFFICIENT IMMUNE RESPONSE 111
4.5 HIGH-ORDERED ASSEMBLIES REPRESENT A NEW PARADIGM OF SIGNAL TRANSDUCTION. 112
5 REFERENCE 113

Abbreviations

1D	One dimensional
2D	Two dimensional
3D	Three dimensional
aa	Amino acid
ASC	Apoptosis-associated speck-like protein containing a CARD
ARIA	Ambiguous restraints for iterative assignment
Å	Ångström
BMRB	Biological magnetic resonance data bank
CARA	Computer-Aided Resonance Assignment
CARD	Caspase activation and recruitment domains
CARDIF	CARD adapter inducing interferon beta
CD	Circular dichroism spectroscopy
CNS	Crystallography & NMR System
COSY	Correlation spectroscopy
CP	Cross-polarization
CPMG	Carr-Purcell-Meiboom-Gill delay
CSA	Chemical Shift Anisotropy
CV	Column volume
DARR	Dipolar assisted rotational resonance
DISC	Death-inducing signaling complex
ddH_2O	Distilled, deionised water
DLS	Dynamic light scattering
DS	Double strand
DTT	Dithiothreitol
E.coli	Escherichia coli
EDTA	Ethylenediamine tetraacetic acid
EM	Electron Microscopy
FID	free induction decay
FTIR	Fourier Transform Infra-Red spectroscopy
HSQC	Heteronuclear single quantum coherence

Hz	Hertz
IFN	Interferon
INEPT	Insensitive nuclear enhancement by polarization transfer
IPS1	Interferon beta promoter stimulator protein 1
IPTG	Isopropyl-β-D-thiogalactopyranoside
K	Kelvin
KPi	KH_2PO_4/K_2HPO_4
LB	Lysogeny broth
MDA5	Melanoma differentiation-associated protein 5
MAS	Magic angle spinning
MAVS	Mitochondrial antiviral-signaling protein
Ni-NTA	Nickel-nitrilotriacetic acid-agarose
NMR	Nuclear Magnetic Resonance
NOE	Nuclear Overhauser effect
NOESY	Nuclear Overhauser effect spectroscopy
OD600	Optical density at 600 nm
PAGE	Polyacrylamide gel electrophoresis
PAIN-CP	Proton assisted insensitive nuclei cross polarization
PAR	Proton Assisted Recoupling
PCR	Polymerase chain reaction
PDB	Protein data bank
PDSD	Proton Driven Spin Diffusion
ppm	Parts per million
PPR	Pathogen pattern recognition receptors
PYD	Pyrin domain
REDOR	Rotational Echo Double Resonance
RF	Radio frequency
RIG-I	Retinoic acid-inducible gene 1
RMSD	Root Mean Square Deviation
RNA	Ribonucleic acid
rpm	Revolutions per minute
SDS	Sodium dodecyl sulphate
SS	Single strand
SSNMR	Solid-state nuclear magnetic resonance

T	Tesla
TEDOR	Transferred Echo Double Resonance
TEM	Transmission electron microscopy
TOCSY	Total correlation spectroscopy
ThT	Thioflavin T
TRIS	Tris(hydroxymethyl)aminomethane
VISA	Virus-induced signaling adapter
wt	Wild type

Summary

Upon virus infection, the innate immune response can be triggered by RIG-I like receptors (RLR) recognizing viral RNA. As a result, the two caspase activation and recruitment domains (CARD)s of RLRs become exposed and activate the CARD of the mitochondrial antiviral signaling protein (MAVS) to form high molecular weight assemblies, which serve as a 'platform' for the activation of downstream signaling components. The N-terminal CARD of MAVS is responsible for this aggregation. Recombinant MAVS CARD forms fibers, which can trigger the conversion of endogenous inactive full-length MAVS into active aggregates. The mechanism of MAVS CARD mediated self-assembly is still poorly understood.

In this thesis, the structure of the MAVS CARD monomer was determined at pH 3.0 by solution NMR. It adopts a typical Greek key bundle of six anti-parallel helices. Analysis of the backbone dynamics revealed the MAVS CARD is extremely rigid, with minor exceptions at the termini. Pull-down assays and NMR titration assays of the RIG-I CARDs with monomeric MAVS CARD mutants provided first insights into the interaction interface between these two proteins. Moreover, solid state NMR was employed to determine the structure of MAVS CARD in its fibrillar conformation. Only minor structural changes were observed relative to the monomeric state. Thus, MAVS activation is one of the first examples of a prion-like conversion from a monomeric to an aggregated state that does not require a substantial conformational change. The quaternary structure of MAVS CARD was determined by combining intermolecular NMR constraints with helical symmetry information from electron microscopy. Three different interaction surfaces were identified between the MAVS CARD domains within the filament. These results presented here will be of significant value for future studies on MAVS activity and regulation, and the role of higher-order assemblies in signal transduction.

Zusammenfassung

Virale RNA wird von Rezeptoren der Rig-I Familie (RLR) erkannt und löst so eine angeborene Immunantwort aus. Die RLRs setzen ihre beiden Caspase-Aktivierungs- und Rekrutierungsdomänen (CARD) frei und aktivieren die CARD des mitochondrialen antiviralen Signalproteins (MAVS). Dieses bildet so hochmolekulare Strukturen, die als Plattform für die Aktivierung weiterer Signaltransduktionsmoleküle dienen. Die N-terminale MAVS CARD ist für diese Aggregation ausreichend. Recombinante MAVS CARD bildet Fasern, die eine Konversion von endogenem, inaktivem Volllängern-MAVS zu aktiven Aggregaten auslösen. Der Mechanismus dieser Selbstassemblierung ist noch schlecht verstanden.

In der vorliegenden Arbeit wurde die Struktur des MAVS CARD Monomers bei pH 3.0 durch Lösungs-NMR bestimmt. Es bildet ein typisches “Greek Key”-Bündel aus sechs antiparallelen Helices. Die Protein-Rückgrat-Dynamik zeigte, das MAVS CARD sehr rigide ist, mit Ausnahme der Termini. Pulldown- und NMR-Tirationsexperimente zur Bindung der RIG-I CARD an monomere MAVS CARD Mutanten haben erste Einblicke in die Interaktionsflächen zwischen den beiden Proteinen gegeben.

Mithilfe von Festkörper-NMR wurde die Struktur von MAVS CARD Fasern bestimmt und gezeigt, dass nur geringfügige Strukturänderungen relativ zum Monomer auftreten. Daher stellt die Aktivierung von MAVS eines der ersten Beispiele einer Prion-ähnlichen Konversion von einem monomeren zu einem aggregierten Zustand dar, der keine substantiellen konformationellen Änderungen beinhaltet. Um die Quartärstruktur von MAVS CARD zu lösen, wurden intermolekulare NMR Distanzinformationen mit Negativkontrast-Elektronenmikroskopie und Röntgenfaserbeugung zur Bestimmung der helikalen Symmetrie kombiniert. Es konnten drei Kontaktflächen der MAVS CARD Protomere untereinander identifiziert werden. Diese Ergebnisse werden für zukünftige Studien zur Aktivierung und Regulation von MAVS, sowie zur Rolle von hochgeordneten Aggregaten in der Signaltransduktion von hohem Wert sein.

1 Introduction

1.1 Pathogen Recognition and Innate immunity

Upon virus infection, host cells mount a remarkable defence comprising the innate and adaptive immune responses which are essential for protecting the host (Iwasaki and Medzhitov, 2010; Takeuchi and Akira, 2009, 2010). Initiation of the innate immune response is mediated by cellular proteins called pathogen pattern recognition receptors (PRRs), which can, for example, distinguish between viral or host-derived nucleic acids. Different PRRs can react with specific viral components knows as pathogen associated molecular patterns (PAMPs), leading to distinct antiviral responses. The basic molecules underlying a monitoring by the innate immune response are highly conserved among species.

1.1.1 Pathogen product recognition by innate immune system

In eukaryotic cells, the pathogen-derived products are recognized by pattern-recognition receptors which can activate signaling cascades, leading to production of interferons (IFNs) and other cytokines. These compounds subsequently mediate innate immune response and modulate adaptive immunity (Katze et al., 2002). There are four categories of PRRs involved in pathogen recognition: Toll-like receptor (TLRs), nucleotide-binding oligomerization domain (NOD)-like receptors (NLRs), RIG-I-like receptors (RLRs) and C-type lectin receptors (CLRs) (Kawai and Akira, 2009; West et al., 2011). CLRs constitute a family of proteins having at least one C-type lectin domain that specifically recognizes carbohydrate structures from pathogens (West et al., 2011). The TLR family members locate on the cell surfaces or the lumen of intracellular vesicles like endosomes or lysosomes, and bind extracellular pathogen components (Uematsu and Akira, 2007). In contrast, the cytosolic pathogen-associated molecules are detected by NLRs and RLRs. NLRs consist of two major subfamilies based on their N-terminal domains, which are caspase activation and recruitment domains (CARD) in the NOD subfamily, or pyrin domains in the NALP (NACHT, LRR and PYD domains-containing protein) subfamily. Some of them are well characterized. NOD1 and NOD2 recognize intracellular bacterial products, such as peptidoglycans, which are cell wall constituents of

both Gram negative and Gram negative bacteria (Strober et al., 2006). NALP3 acts as a regulatory domain responding to multiple stimuli to form a complex termed the NALP3 inflammasome, which promotes caspase 1 dependent release of IL-1ß (Subramanian et al., 2013). RLRs belong to a group of RNA helicases that specifically detects virus derived RNA species. They recruit downstream factors via CARD to initiate anti-viral response through type I IFN induction. Details on the mechanism of nucleic acid recognition by PRRs will be discussed in the following sections.

1.1.2 Toll like receptors involved in the detection of nucleic acids

TLRs play an important role in the detection of viral infections. They are a family of 10 proteins in human and consist of leucine-rich repeats (LRRs), a transmembrane domain, and a conserved cytoplasmic signaling domain designated the Toll/IL-1receptor homology (TIR) domain (Akira et al., 2006). Four members in this family are implicated in nucleic acid detection, TLR3, TLR7, TLR8, and TLR9 (Takeuchi and Akira, 2009). These receptors are located on membranes of cytoplasmic vesicles such as endosomes, lysomes and the endoplasmic reticulum (ER), where they monitor the lumen of these cellular organelles for microbial nucleotides. TLR3 recognizes dsRNA with a minimum length of 40 base pairs in conventional dendritic cells (Botos et al., 2009), while TLR7/8 and TLR9 recognize guanosine-uridine rich ssRNA and DNA with nonmethylated cytosine-guanosine (CpG) motifs, respectively (Heil et al., 2004). Activation of these receptors leads to induction of type I interferon, production of inflammatory cytokines and upregulation of major histocompatibility complex (MHC) to mediate antiviral responses.

1.1.3 RIG-I like receptors and their regulation in antiviral immunity

While the TLRs' recognition of extracellular pathogen-associated nucleic acid, RLRs are responsible for the detection of virus-associated nucleic acids present in the cytosol. As this is typically the case for directly infected cells, RLRs are considered to act as more cell-autonomous sensors for viral infections than TLRs. This notion is supported by the fact that RLRs are broadly expressed in many different cell types (Barbalat et al., 2011; Kato et al., 2006). RLRs recognize viral replication through direct interaction with short double strand or 5´ ends phosphated single strand viral RNA (Barbalat et al., 2011). These are typically not present in uninfected host cells, because host mRNAs are capped

with a 7-methyl-guanosine group, tRNAs are further processed, and mRNAs associate with ribosomal proteins.

1.1.3.1 Members of RLR family

Three members of the RLR family have been found in mammals: retinoic acid-inducible gene 1 (RIG-I) and melanoma differentiation associated gene 5 (MDA5), which are responsible for the recognition of viral RNA and for the activation of antiviral signaling, and laboratory of genetics and physiology 2 (LGP2), which appears to serve as a dominant negative inhibitor of RNA virus induced responses (Yoneyama and Fujita, 2008) (**Figure 1**). RIG-I and MDA5 each contains two N-terminal caspase-recruitment domains (CARDs), a central DExD/H box RNA helicase/ATPase domain, and a C-terminal regulatory domain (RD), whereas LGP2 lacks the CARD domains and consists of the DExD/H box helicase/ATPase domain, and a C-terminal domain. The DExD/H box helicase domain

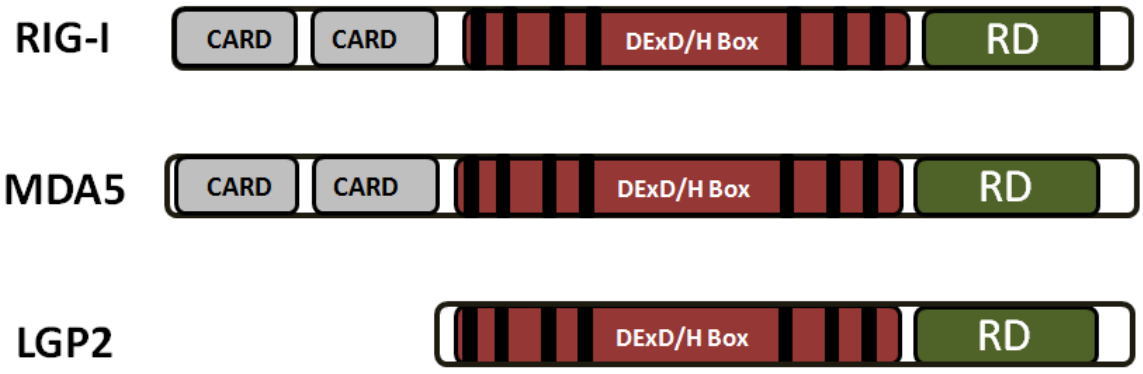

Figure 1. Structural domains of members of RLRs family: Sequencing alignment reveals that LGP2, MDA5 are homologous to RIG-I with around 30% identity in the amino acid sequence. All three members are evolutionary conserved among vertebrates (Zou et al., 2009), suggesting their critical role in innate immunity.

is involved in RNA binding and its ATPase function is important for driving a conformational change to initiate the signaling transduction (Saito and Gale, 2007). The tandem CARD domains facilitate their interactions with other proteins via CARD-CARD interactions, which trigger the downstream signaling. The C terminal regulatory domain interacts with CARD to sequester it in

unstimulated states to keep it inactive. For this reason, it is also called C terminal repressor domain. Since LGP2 lacks a CARD, it suggests that LGP2 can not initiate downstream signaling. The in vivo experimental data supports the conclusion that LGP2 functions as a negative regulator of the RIG-I pathways (Yoneyama et al., 2005).

1.1.3.2 Distinct recognition pattern of RLRs

RIG-I and MDA5 have distinct and divergent recognition patterns for viral RNA (**Figure 2**). While 5´ triphosphate-group containing RNA, and short double strand RNA are detected by RIG-I, longer than 2kbp double strand RNA is preferentially recognized by MDA5 (Barbalat et al., 2011). Recent structural studies showed that MDA5 binds the stem of long dsRNA, while RIG-I binds the end of short dsRNA (Peisley et al., 2013; Wu et al., 2013). The structure of their complex with dsRNA revealed that different flexibility of a loop in the C terminal domain (CTD) could cause a different

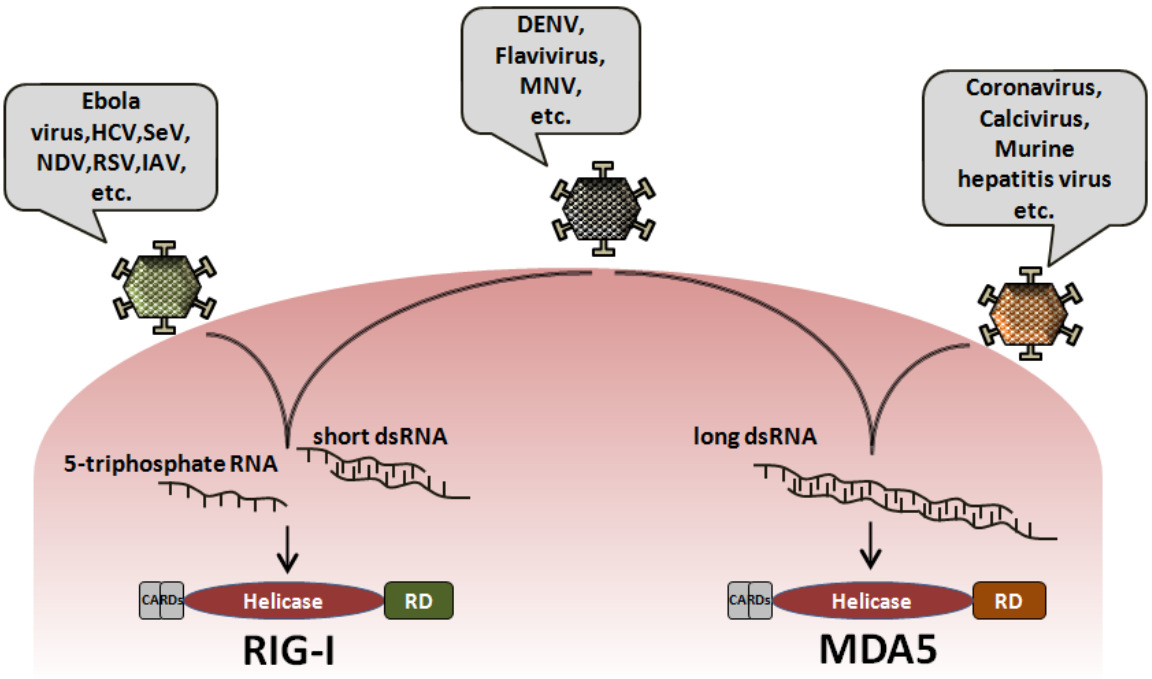

Figure 2. Distinct recognition pattern of RIG-I and MDA5 and their role in virus detection: MDA5 binds with long dsRNA, while RIG-I discriminates 5'-triphosphate RNA and short dsRNA. The virus detected by MDA5 and RIG-I are delineated by arrows. Some viruses can activate both at the same time.

orientation of the CTD. As a result, the dsRNA is bound either in a open C ring or a closed O ring. In RIG-I the loop rigidly binds the end of dsRNA. In MDA5 the loop provides a docking site to hold the dsRNA stem. Interestingly, both RIG-I and MDA5 form signaling competent filaments. Comparing to the monomer, the filaments greatly enhance downstream MAVS filament formation. Many studies showed that RIG-I and MDA5 are involved in the recognition of different viruses. Japanese encephalitis virus, hepatitis C virus (HCV) (Kato et al., 2006; Loo et al., 2008), Ebola virus, Filovirus (Cardenas et al., 2006), Vesicular stomatitis virus (VSV) and Rabies virus (Faul et al., 2010), Orthomyxoviruses, Influenza A,B virus (IAV, IBV); Sendai virus (SeV), Newcastle disease virus (NDV), and Respiratory syncytial virus (RSV) (Kato et al., 2006) are found to be detected by RIG-I. Whereas, Murine hepatitis virus, Coronavirus, Murine norovirus-1 and the Calicivirus (Kato et al., 2006; Yoneyama and Fujita, 2009) are recognized by MDA5. Some viruses are able to active both MDA5 and RIG-I at the same time such as West Nile virus (WNV), Dengue virus (DENV) Flaviviruses, Rotavirus as well as Reovirus (Broquet et al., 2011; Loo et al., 2008).

1.1.3.3 Signal pathway of RLR mediated immune response

Upon activation by viral nucleic acids, RIG-I and MDA5 associate with the adapter protein mitochondrial antiviral signaling (MAVS), which is also named virus-induced signaling adapter (VISA), or CARD adapter inducing IFN-b (CARDIF) (Kawai et al., 2005; Meylan et al., 2005; Seth et al., 2005; Xu et al., 2005). MAVS also contains an N-terminal CARD domain through which it exhibits interaction with RIG-I and MDA5. This interaction results in the assembly of MAVS into high molecular weight structures (Hou et al., 2011; Tang and Wang, 2009). It has been suggested that K63-linked ubiquitin chains are involved in the catalytic activation of MAVS assembly by RIG-I (Zeng et al., 2010). The assembled MAVS then serves as a scaffold mediating its interaction with tumor necrosis factor receptor-associated factors (TRAF2, TRAF5, TRAF3 and TRAF6), receptor interacting protein 1 (RIP1) and Fas-associated death domain (FADD) (Kawai and Akira, 2009; Liu et al., 2011; Meylan et al., 2005; Michallet et al., 2008; Seth et al., 2006). This orchestration leads to the activation of interferon regulatory factor 3 (IRF-3) and nuclear factor κB (NF-κB). These two factors then move into the nucleus to initiate the expression of type I interferon and other inflammatory cytokines (**Figure 3**). In addition, another molecular stimulator of interferon genes (STING) also known as MITA was revealed to interact with MAVS to potentiate IRF and (NF-κB) activation (Ishikawa and Barber, 2008; Zhong et al., 2008). STING is a transmembrane protein from

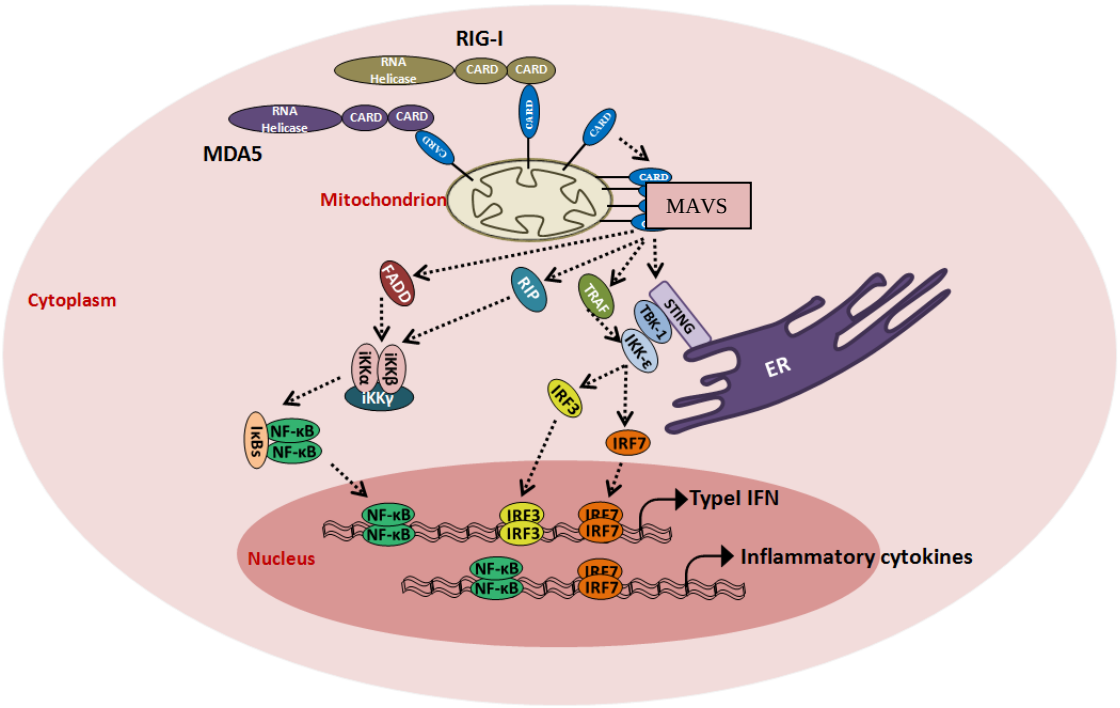

Figure 3. RLR signaling pathway induced by virus infection: RIG-I and MDA5 mediate signaling through recruitment of MAVS via CARD-CARD interactions. MAVS in turn activates IRF3/7 and NF-κB by the kinases TBK-1, IKK-ε, RIP, FADD and STING leading to expression of type I interferon and other proinflammatory cytokines. STING is located on the membrane of the ER, the interaction of it with MAVS showing the importance of ER–mitochondrial tethering after virus infection. LGP2 works as a negative regulator of RIG-I and MDA5 pathway probably to balance between antiviral immune response and autoimmune response.

the endoplasmic reticulum (ER). This discovery highlights the importance of interconnectedness between mitochondria and ER in the innate immune response.

1.1.3.4 Regulation of RLR signaling cascade

RIG-I mediated signaling is revealed to be positively and negatively controlled by ubiquitination of RIG-I CARDs by tripartitie motif 25 (TRIM25) and ubiquitin ligase RNF125 respectively. TRIM 25 is a ubiquitin E3 ligase which is important for efficient activation of RIG-I. The TRIM25$^{-/-}$ cells display reduced production of type I interferon upon virus infection (Gack et al., 2007). On the other hand, a recent paper shows that ATP-driven formation of RIG-I filaments around the dsRNA

substrate is sufficient to enhance the CARD-driven activation of MAVS, without cofactors like K63 ubiquitins (Peisley et al., 2013). However, RNF125 mediated modification of RIG-I with ubiquitin leads to its degradation by the proteasome (Arimoto et al., 2007). RIG-I signaling could also be regulated through its direct interaction with inhibiting and activating factors. The proteasome subunit alpha type7 (PSMA7) was revealed to bind MAVS and impair its ability to translate signaling (Jia et al., 2009). Poly (rC)-binding protein 2 (PCBP2) was also found to mediate MAVS ubiquitination and degradation via E3 ubiquitin ligase atrophin-interacting protein 4 (You et al., 2009). Autophagy conjugate (ATG5 and ATG12) was shown to interact with both RIG-I and MAVS, inhibiting their ability to produce IFN (Jounai et al., 2007). In contrast, the protein zinc finger antiviral protein short isoform (ZAPS) interacts with RIG-I to promote the signaling actions (Hayakawa et al., 2011).

1.1.4 MAVS – an essential adaptor in RLR antiviral pathway

MAVS consists of 540 amino acids that form an N-terminal CARD domain, a proline-rich domain and a transmembrane domain (**Figure 4 A**). The region between CARD and the transmembrane domain including the proline rich domain is predicted to be random coil (**Figure 4 C**). This long random coil domain contains the binding sites for many downstream signaling proteins such as TRAF, RIP1 and FADD (Michallet et al., 2008). The transmembrane domain locates MAVS on the outer membrane of mitochondrion in a variety of cell types. In addition, MAVS can also be positioned at the membrane of peroxisomes, where it is involved in the early induction of interferon stimulated genes (ISGs), but not type I interferon (Dixit et al., 2010). A construct comprising the CARD domain and the transmembrane domain, the so called mini-MAVS, seems to be the minimal requirement for its cellular function, because over expression of mini-MAVS is sufficient for signal transduction (Seth et al., 2005). Alignment of the CARD sequence among different species revealed that 70% of the sequence is identical during evolution, indicating the critical role it has for cells (**Figure 4 B**). MAVS activation is essential to innate immune signaling. Studies on MAVS knockout cells demonstrated that it was required for viral induction of interferon and cytokines in multiple cell types, including dendritic cells, fibroblasts and macrophages (Barbalat et al., 2011). Furthermore, MAVS deficient mice failed to induce an immune response upon stimulation by SeV and poly(I:C) (Sun et al., 2006).

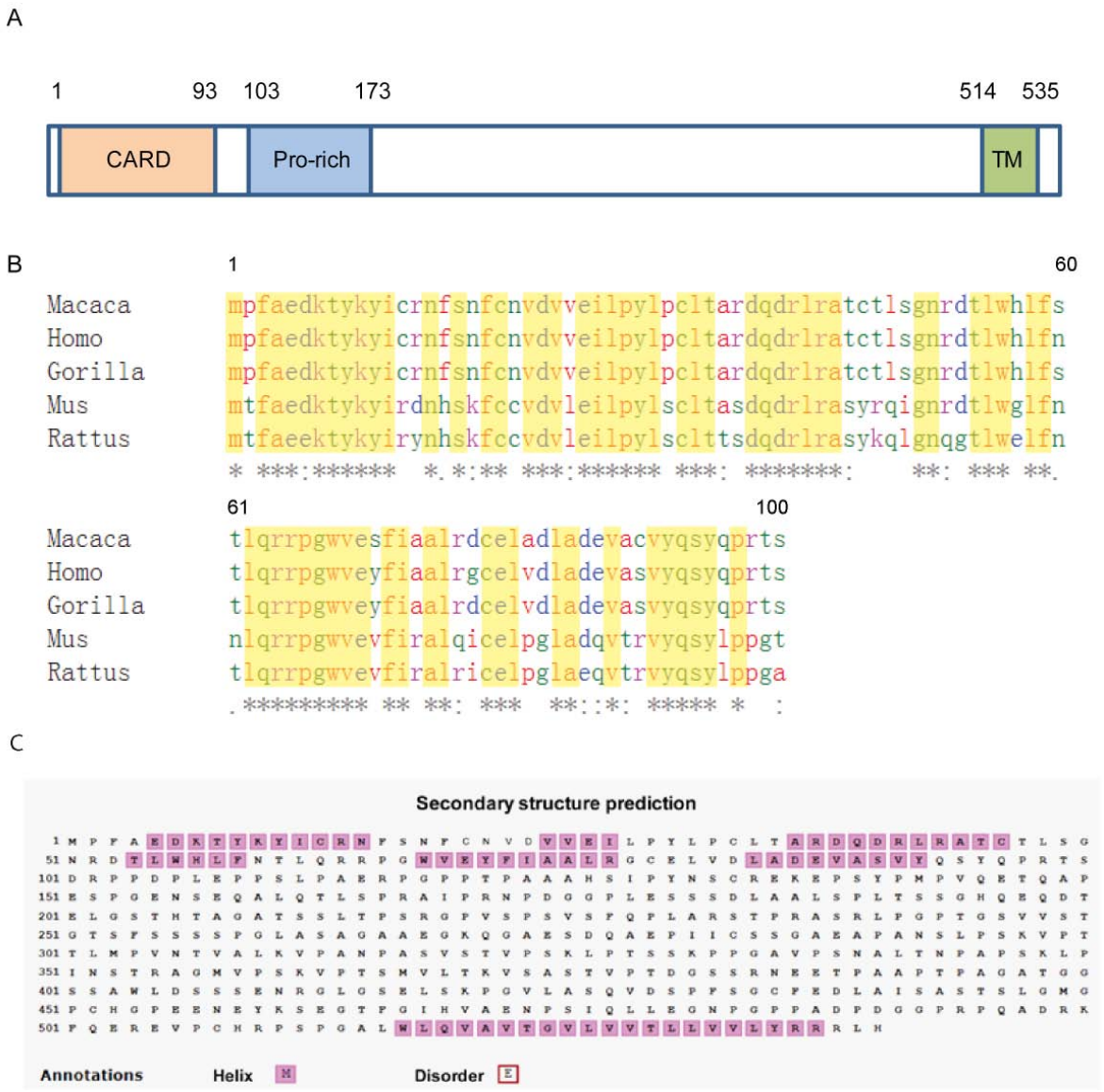

Figure 4. Secondary structure of MAVS and sequential alignment of its CARD between different species: A. domain organisation of MAVS. **B.** sequence alignment of MAVS CARD from different mammalian species. The CARD domain are highly conserved, ~70% of the residues are the same among different mammalian species, indicating a critical role during evolution. **C.** Secondary structure prediction of full length MAVS using Jpred3 (Cole et al., 2008). Helix region is labeled in pink and random coil region is labeled in write.

1.1.4.1 MAVS in type I interferon production and inflammasome activation

Upon activation by RIG-I, MAVS with a predicted molecular weight of 56kD is assembled into a supramolecular complex in size larger than the 26s proteasome (Hou et al., 2011; Moresco et al., 2011). After MAVS has formed signalling-competent oligomer, the proline rich region interacts

with TRAF family members including TRAF2, TRAF3, TRAF5 and TRAF6. These E3 ubiquitin ligases catalyse K63 polyubiquitination of TRAF2 and other target proteins, which are then bound by the ubiquitin-binding protein NEMO. The TANK/TBK1 and IKK complexes are then recruited by NEMO and subsequently activated, resulting in the phosphorylation of IRF3 and IkBa, (Liu et al., 2011; Saha et al., 2006; Xu et al., 2005).

It was shown that during activation of MAVS, its dimerization through the transmembrane domain is critical for inducing interferon (Baril et al., 2009). However, other studies showed that the transmembrane domain can be replaced by other mitochondrial proteins (Tang and Wang, 2009) and oligomerization of MAVS in the cytosol is sufficient to induce the type I interferon production independent of the mitochondrial localization (Takamatsu et al., 2013). MAVS-dependent innate antiviral immune response occurs from both mitochondria and peroxisomes with different kinetics. Mitochondrial MAVS mediated interferon-dependent, delayed immune signaling for a long-term containment of the infection, while peroxisomal MAVS induces immediate, interferon-independent expression of antiviral factors for the nascent infection (Dixit et al., 2010).

MAVS has also been found to associate with NLRP3 and promote its oligomerization leading to inflammasome activation. Recent studies from two groups revealed that MAVS mediates recruitment of NLRP3 to mitochondria, inducing NLRP3-dependent caspase-1 activation and production of IL-1β (Park et al., 2013; Subramanian et al., 2013). These new findings show that in addition to its well-studied role in antiviral immunity, MAVS is also a critical adaptor required for inflammasome function, which further supports a role of mitochondria as a platform integrating multiple innate immune events.

1.1.4.2 Prion-like behaviour of MAVS in RIG-I signaling

Recently, the group of Zhijian Chen used a cell free system reconstituting RIG-I mediated innate immune signaling pathway to show that MAVS forms functional prion-like fibrillar aggregates to propagate RIG-I signaling (Hou et al., 2011). They found that both *in-vitro* formed recombinant MAVS polymers, and *in vivo* activated MAVS aggregates could induce assembly of endogenous MAVS on mitochondria in the presence of cytosolic extracts. In addition, they showed that this prion-like capacity of MAVS converting endogenous MAVS into an aggregated form was solely dependent on its CARD domain. Thus they established a new paradigm that innate immune signaling could be amplified *via* an autocatalytical conformational conversion of MAVS. The prion

phenomenon, i.e. a transmissible conformational change of protein structure, has been first described for the prion protein (PrP), where the prion conformation is replicated through contact between endogenous properly folded PrP^{c} (**C**ommon) with disease linked, misfoled PrP^{sc} (**Sc**rapie). The latter PrP^{sc} converted PrPc from a mostly α-helical structure to a β sheet-rich aggregated form. There are some similarities between MAVS and PrP^{sc}. First, both have the ability to convert unaffected molecules into aggregated fibrillar form which is stable and resistant to protease and detergent. Second, the newly induced MAVS or PrPsc aggregates can act as seeds to convert more endogenous molecules into the aggregated form. Despite these similarities, MAVS aggregates are functional and its aggregates cannot be stained with Congo red suggesting it does not have a β sheet rich structure. To understanding the mechanism of MAVS aggregation *in vivo*, many questions remain to be answered such as: does MAVS forms fiber *in vivo*? how is the aggregation initiated and how it is resolved or degraded afterwards? How does the nucleation occur on the surface of mitochondria? As mentioned above, the RNA-bound filaments of RIG-I and MDA5 trigger the activation of MAVS. This process resembles the *de novo* formation PrP^{sc}, where PrP^{sc}-like oligomers are believed to act as nuclei for the aggregation process (Biasini et al., 2012). RIG-I and MDA5 could thus be considered to be the nucleators for MAVS assembly into its aggregated, functional form.

1.1.4.3 Regulation of MAVS by mitochondrial membrane and proteins

Insufficient or enhanced levels of immune response can be detrimental, leading to the failure to remove the invading threat or causing cell damage, respectively. Tight regulation is thus a prerequisite to control antiviral signaling. Mitochondria represent a critical platform converging signals propagated from different cytosolic antiviral sensors. Therefore, MAVS is positioned at a strategically important position of the antiviral innate signalling cascade. Till now several different ways of modulating MAVS expression and signaling have been reported (**Figure 5**).

1.1.4.3.1 Regulation through protein-protein interaction

Mitochondrially localized proteins are logical candidates for interaction with MAVS. For instance, the protein NLR family member X1 (NLRX1), which is located on the outer mitochondrial membrane, interacts with the CARD domain of MAVS, disrupting RIG-I activation of MAVS (Saha

et al., 2006). An NLRX1 deficient cell line has increased IFNβ production. However, conflicting results exist. Two other groups found no potentiation of IRF3 phosphorylation and IFNβ upregulation in NLRX1 knock out cell lines and mice upon stimulation by SeV or poly(I:C) (Rebsamen et al., 2011; Soares et al., 2013). The mitofusins (MFN1 and MFN2), which are effectors

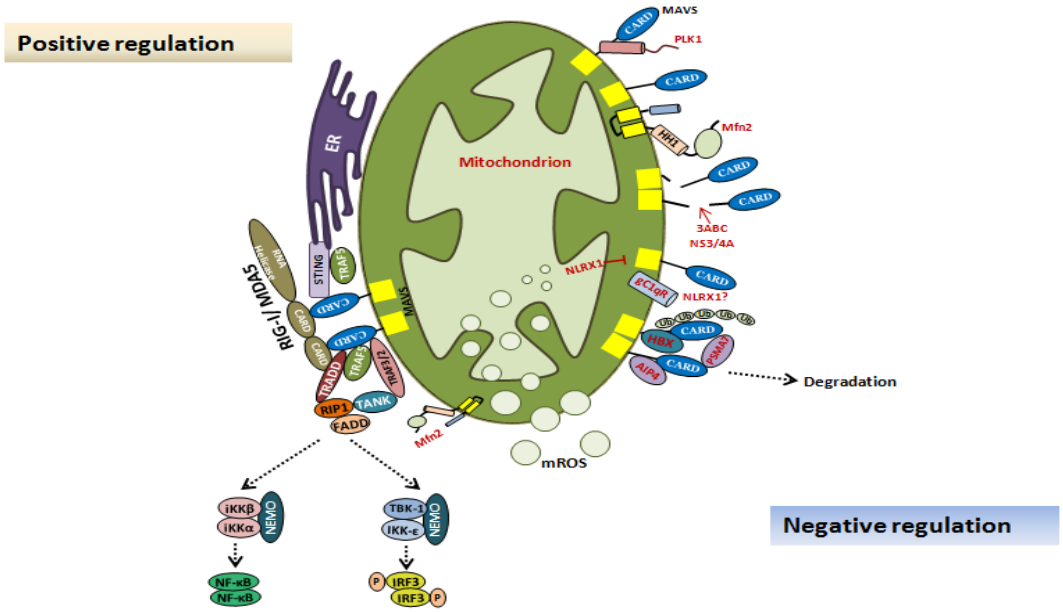

Figure 5. Regulations of MAVS-mediated antiviral immune response: MAVS is activated by virus infection. Signaling through multiple proteins leads to dimerization of IRF3 and activation of NF-κB which transfer into the nucleus initiating the expression of IFNs. The signal pathway is depended on mitochondrial membrane potential and also regulated by mitochondrial ROS. Increasing mROS induces upregulation of MAVS signaling. MAVS also interacts with many mitochondria-localized proteins for both positive and negative regulation. Viral proteins also interact with MAVS in order to antagonise RLR signaling. In addition, post-translation modifications such as ubiquitination and phosphorylation modulate MAVS activity.

of mitochondrial fusion events, are also located on the outer mitochondrial membrane. MFN2 were reported to bind and interact with MAVS (Yasukawa et al., 2009). Further investigations showed that MFN2 hydrophobic repeat (HR1) region bound to the C-terminal region of MAVS, and this interaction inhibited MAVS mediated IFN production (Yasukawa et al., 2009). The receptor of globular head domain of complement component C1q (gC1qR) is a protein that localizes to mitochondria when viruses infect a cell. gC1qR was shown to inhibit the RLR signaling pathway

(Xu et al., 2009). While the interaction between gC1qR and MAVS was weak in uninfected cells, it was greatly enhanced by viral infection (Xu et al., 2009).

In contrast to these negative regulators, there are also several positive regulators of MAVS. The resident mitochondrial translocase of outer membrane protein (Tom 70) was identified to be involved in the complex of IRF3/TBK1 and MAVS to facilitate signaling. A proposed model was Tom 70 interacted with Hsp90/IRF3/TBK1 complex and recruited them to mitochondria promoting MAVS interaction with its downstream targets (Yang et al., 2006; Young et al., 2003). Another study on IFN-induced protein with tetratricopeptide repeats 3 (IFIT3) showed that IFIT3 localized to mitochondria and served as a scaffold to facilitate MAVS interaction with TBK1. Overexpression of IFIT3 increased the RLR signaling and IFN production. On the other hand, the depletion of it attenuated the immune response (Liu et al., 2011).

1.1.4.3.2 Regulation through mitochondria dynamic and membrane potential

Mitochodrial dynamics including homotypic fusion and fission, as well as changes in mitochondrial morphology play an important role in fundamental functiona of mitochondria such as maintainance of mtDNA and apoptosis (Suen et al., 2008). Recent studies demonstrated that physical properties of the mitochondria also have an effect on antiviral immunity by changing MAVS distribution and aggregation states. Arnoult and colleagues reported that infection of cells by SeV, leading to RLR signalling, also caused an elongation or fusion of mitochondria (Arnoult et al., 2009). They also demonstrated that MAVS interacted with Mfn1 to positively regulate MAVS mediated immune response (see 1.1.4.3.1). Interestingly, when they knocked down Mfn1 or optic atrophy 1 (OPA1), the production of IFN-β was decreased. Conversely, knocking down of mitochondrial fission factors, dynamin-related protein 1 (Drp1) and Fis-1, increased the production IFN-β. Thus, they proposed that the dynamics of mitochondria enhances the association of MAVS with STING, another antiviral signaling adaptor localized on ER membrane, and facilitates MAVS downstream signaling (Castanier et al., 2010).

Further evidence implied that the mitochondrial membrane potential ($\Delta\psi_m$) and mitochondrial reactive oxygen species (mROS), also played an important role in MAVS mediated innate antiviral immunity (Koshiba et al., 2011). Two recent studies reported that the mitochondrial membrane potential is essential for MAVS-mediated signalling: Cells treated with the protonophore CCCP, dissipating the membrane potential $\Delta\psi m$, are defective in antiviral immune response (Koshiba et al.,

2011). Secondly, overexpression of uncoupling protein-2, which induced mitochondrial proton leakage resulting in $\Delta\psi_m$ dissipation, greatly suppressed the innate immune response (Krauss et al., 2005). In addition, mitochondrial ROS were also reported to be involved in the RLR signaling pathway. Cells depleted of the mitochondrial cytochrome c oxidase 5B (COX5B), which is responsible for repression of ROS, exhibited increased antiviral immune response (Zhao et al., 2012). All of the above studies provide a frame for understanding mitochondria's role in innate immunity.

1.1.4.3.3 Regulation through ubiquitination and phosphorylation of MAVS

Many proteins are regulated by post-translation modifications to modulate the half life and other aspects of them. For MAVS, it was also identified to interact with proteasome subunit alpha type-7 (PSMA7), which induced the ubiquitination of MAVS, facilitating the proteasomal degradation of MAVS (Jia et al., 2009). PSMA7 is part of ring-shaped 20S core structure of the proteasome. It consists of 7 alpha subunits and is involved in the regulation of proteasomal activity. Overexpression of PSMA7 reduced MAVS protein level and suppressed the MAVS mediated antiviral signaling, Another study by You and colleagues, identified poly(rC) binding protein 2 (PCBP2) as a regulator mediating MAVS degradation. They identified MAVS C-terminal amino acids 460-540 to be involved in binding with PCBP2. Later they found that Nedd4-like E3 ubiquitin ligase (AIP4) interacted with MAVS in presence of PCBP2, suggesting PCBP2 worked as scaffold for ubiquitination of MAVS by AIP4. Mutagenesis experiments revealed that K371 and K420 of MAVS are the two critical points for proteasome dependent degradation (You et al., 2009). TRIM25 mediated ubiquitination of MAVS was also reported to play a role (Gack et al., 2007). It enhanced production of IFN-β after stimulation by poly(I:C). Depletion of TRIM25 led to attenuation of immune response. Conflicting result by Castanier and colleagues showed that overexpression of TRIM25 promoted the ubiquitination and degradation of MAVS. They identified the ubiquitination sites on MAVS were K7 and K10 (Castanier et al., 2012). It remains to be seen the role of TRIM25 in regulation of MAVS. In addition to ubiquitination, phosphorylation is also found to be involved in regulation of MAVS. Yeast two hybrid assay identified Polo-like kinase 1 (PLK1), a serine and threonine kinase phosphorylating MAVS at position Ser233 and Thr234. The phosphorylation attenuated IFN signaling by disrupting one of the TRAF3 binding sites (Vitour et al., 2009). More recently, the tyrosine kinase c-Abl was identified to be another interaction partner of MAVS. It

acted as a positive regulator to increase MAVS mediated IFN signaling (Song et al., 2010). A tyrosine-scanning mutational analysis revealed that Tyr 9 is required for IFN signaling. The phosphorylation on Tyr9 started to increase 15 min post-infection and reach maximum at 60 min post-infection. Later the level of phosphorylation decreased. Y9F mutant showed impaired antiviral immune response suggesting Y9 phosphorylation might play a positive regulation role (Wen et al., 2012). It is still unknown if c-Abl is related to phosphorylation of MAVS Tyr9.

1.1.4.4 Viral antagonism of MAVS mediated signalling

In order to escape from the host cell immune response, pathogenic viruses have developed mechanisms to disrupt RLR signaling. Many viral proteins have been identified as antagonists to block MAVS mediated antiviral immune response (**Figure 6**).

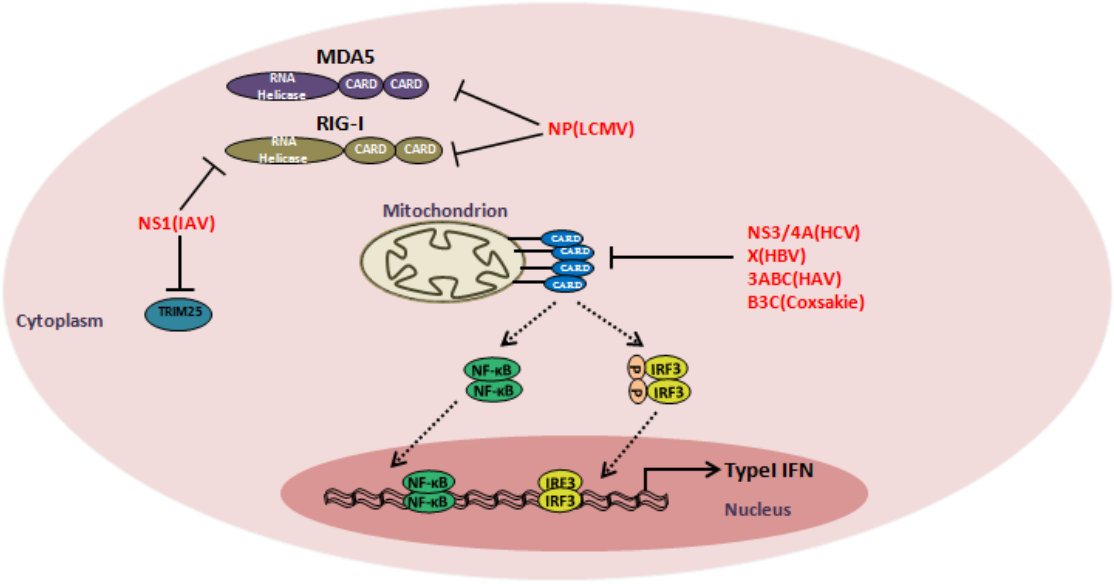

Figure 6. Viral antagonism of RLR signaling: Many viruses have developed mechanisms to disrupt RLR signaling pathway in order to escape the innate immune response. Numerous viral proteins target and bind components of RLR signalosome. Some known viral proteins are shown with lines to their targeted proteins.

Several viruses encode proteases to cleave host cellular targets to suppress antiviral immunity. Hepatitis C virus serine protease NS3/4A cleaves MAVS at position C508, removing it from

mitochondria and abolishing its downstream signaling pathway (Li et al., 2005). Hepatitis A virus express cystine protease precursor 3ABC to cleave MAVS in a similar way (Yang et al., 2007). In addition, the Coxsackie virus was also identified to express protease B3C to cleave MAVS (Mukherjee et al., 2011). Other viruses can block the RLR signaling pathway by sequestering signaling components or targeting them for proteasome mediated degradation. Influenza A virus protein NS1 was shown to sequester RIG-I and TRIM25 from their signaling complexes to reduce the stimulation of downstream MAVS (Gack et al., 2009; Mibayashi et al., 2007). The nuclear protein (NP) of lymphocytic choriomeningitis virus (LCMV) was shown to interact with both RIG-I and MDA5 to reduce the production of IFN (Zhou et al., 2010). X protein of Hepatitis B virus could target MAVS for degradation through K63 ubiquitination (Wei et al., 2010). Vpr, Vif protein of human immunodeficiency virus (HIV) were found to directly deplete MAVS downstream signaling component IRF-3 to block the production of IFN (Okumura et al., 2008). PB2 protein of Influenza A virus was identified to interact with MAVS residues 1-150 via its residues 1-37, especially the helix3 which was observed to promote interaction with MAVS. Interestingly, MAVS bound to the same site on PB2 of MAVS as the PB1 protein binding domain, indicating the importance of this region of PB2 (Patel et al., 2013). The binding of PB2 with MAVS inhibit the expression of IFN-beta. Another protein from Influenza A virus, PB1-F2, was also identified to interact with MAVS, blocking the antiviral signaling pathway. It was shown that PB1-F2 protein caused a decrease of the mitochondrial membrane potential (Varga et al., 2012).

1.1.5 Structural studies on CARD and other DD members

CARD domains are a subfamily of the death domain (DD) superfamily. Their structures are quite similar to other DD superfamily members such as pyrin domain (PYD), death effector domain (DED), and death domain (DD). They all have typical six helices in Greek key topology. The DD family members are usually involved in the assembly of signaling complexes either by self-association or by interaction with domains from the same subfamily. In the human genome, there are 32 DDs, 7 DEDs, 28 CARDs and 19 PYDs (Kohl and Grutter, 2004; Reed et al., 2004). Although all members of the DD superfamily have conserved structural features, each subfamilies also exhibits distinct structural and surface characteristics. The structures of CARDs is unique for helix 1 which tends to be either bent, to have a kink or to be broken into two helices (Qin et al., 1999; Vaughn et al., 1999). In addition, the surfaces of CARDs are characterized by having positively and negatively

charged patches that might be involved in association. The first complex of CARD-CARD was the heterodimer between Caspase-9 with apoptotic protease activating factor 1 (Apaf-1). It is formed by positively charged helix 1 and helix 4 with negatively charged helix 2 and helix 3 (Qin et al., 1999). This type of interaction later was known as type I interaction. Shortly after the report of Apaf1/Caspase 9 complex, another heterodimer comprising the DDs of Tube and Pelle was published (Xiao et al., 1999). Here, two DDs interact through the helix 4- helix 5 loop with the helix 5- helix 6 loop. There are also electrostatic interaction between basic and acidic amino acids involved in this type of interaction. Finally, the structure of the PIDDosome revealed a third type of interaction formed by the interaction of helix 3 with helix1- helix 2 and helix 3- helix 4 loops (Park et al., 2007b). Later, more complexes of DDs were published showing that DDs associated into high molecular weight helical structures through three types of interactions (**Table 1**) (Lin et al., 2010; Lu et al., 2014; Qiao et al., 2013; Wang et al., 2010). The asymmetric three types of interaction currently present the most common mechanism of association of DDs into polymers. The formation of helical

Table 1. Complexes formed by DDs

DD assemblies	Structure and function
DISC	A complex consists of 5 Fas and 5 FADD DDs through 3 type interactions. It transduces a downstream signaling cascade resulting in apoptosis (Wang et al., 2010).
Myddosome	A complex consists of 6 MyD88, 4 IRAK4, and 4 IRAK2 DDs through 3 types of interactions. It is required for signaling of toll-like receptors in innate immune response (Lin et al., 2010).
$Pyrin^{ASC}$ filament	Apoptosis-associated speck-like protein (ASC) pyrin domains assemble into fibrillar aggregates through 3 types of interactions. It bridges the sensor proteins and caspase-1 to form inflammasome complexes through pyrin domains (Lu et al., 2014).

structures has several biological benefits. It is relatively simple to assemble and to disassociate, and it allows for interactions between multiple partners. Secondly, the helical polymers could work as a platform to increase local concentration of target proteins, facilitating the weak binding for signal propagation. However, the mechanism of how a particular CARD domain distinguishes its interacting CARD partners from other CARDs is still unclear. More studies, also from a structural view, need to be carried out.

1.1.6 High-order assemblies as platform for signal transduction

Recently, several studies showed that proteins form functional high-order assemblies, representing a new paradigm of signal transduction. For example, the so-called PIDDosome complex, which consists of 12 DDs (7 RAIDD DDs and 5 PIDD DDs), forms a double-stranded left-handed helix. Within each strand, the monomers are connected by type III interaction, while type I and II interactions are responsible for inter-strand connections (Park et al., 2007b). The PIDDosome could not only active caspase-2 through CARD-CARD association to initiate apoptosis (Tinel and Tschopp, 2004), but also interact with DD of RIP 1 to active NF-kB (Janssens et al., 2005). The PIDDosome serves as a platform controlling cell life or apoptosis. The structure of PIDDosome explained how DDs formed helical complex. It was a start point to study mechanism of DDs forming filaments, such as death effector filament formed by DEDs from caspase-8/FADD (Siegel et al., 1998) and innate immunity adaptor filament formed by MAVS (Hou et al., 2011). In addition to helical assemblies of DDs, another two ways of forming high-ordered signaling complex were identified. One is RIP1/RIP3 necrosome forming a functional amyloid signalsome. The fibrils have typical character of β-amyloids shown as binding of Thioflavin T and Congo Red. It is assembled through stacking of β strands to form a long β sheet, which is essential for the TNFR induced necrosis (Li et al., 2012). The other is the complex formed by TRAF6, the dimerization of the N terminal and trimerization of the C terminal provided a way for 2D elongation of the oligomers facilitating signal transduction of many immune response (Yin et al., 2009). In conclusion, proteins in cells might work in a way similar to a complicated machine where large protein assemblies serve as platforms or scaffold for many interacted proteins especially enhancing the weak binding between two partners.

1.2 Nuclear Magnetic Resonance (NMR) Spectroscopy

The aim of this part is to explain basic concepts of Nuclear Magnetic Resonance (NMR) that will appear in this thesis. NMR is a physical phenomenon that nuclei can absorb and emit electromagnetic radiation in a magnetic field. The energy should be at specific resonance frequencies which depend on properties of the nucleus and the strength of magnetic field. Many techniques exploit NMR phenomena to study protein structure and dynamics.

1.2.1 Elementary principles of NMR spectroscopy

All atomic nuclei consist of nucleons that are neutrons and protons. The atomic nuclei with odd numbers of nucleons have the intrinsic quantum property of spin (Cavanagh et al., 2007; Evans, 1995). The nuclei with even numbers might or might not have this property. A spin is characterized with a magnetic moment (μ) via the relation

$$\mu = \gamma I, \quad \mu z = \gamma I z \tag{1.1}$$

where γ is the gyromagnetic ratio and I is the nuclear spin angular momentum. Iz and μz mean z component of I and μ

The nucleus spin in other words is a charged and spinning particle, well known to produce a magnetic field itself. When placed in an external magnetic field, the interaction between the nuclear magnetic moment and the external magnetic field makes different nuclear spin states have different energies states (Cavanagh et al., 2007). This difference results in a slightly increased population bias towards the lower energy state, resulting in a net magnetization M_0 parallel to the external magnetic field. Transition between these states will occur only when the nuclear spins absorb or emit specific energy in the radio frequency range. The energy is given by

$$\Delta E = \gamma h B_0 \tag{1.2}$$

h is Planck's constant and B_0 is the strength of external magnetic field.

In a classical view, the nuclei in an external magnetic field B_0 will rotate around the axis of the external field vector in a way called Larmor precession. The frequency of this precession v_0 is the Larmor frequency which equals the transition radio frequency described above. In order to make nuclear magnetization observable, we apply a radio frequency pulse perpendicular to the external magnetic field. The nuclei with a Larmor frequency identical to that of the radio frequency will resonate, causing the net magnetization M_0 to rotate around the direction of this radio frequency pulse. After a pulse long enough to cause a 90° rotation, the nuclear magnetization vector $M_{x,y}$ will rotate in the plane perpendicular to the magnetic field B_0 and decrease with time (Jardetsky and

Roberts, 1981; Wüthrich, 1986). The rotating magnetic field caused by $M_{x,y}$ can be detected via the receiver coil by inducing an electric current. The declining amplitude with time is referred to as the free induction decay (FID), which can be converted into the frequency domain through Fourier transformation (**Figure 7**). This is the way that NMR spectra are recorded and displayed (Ernst et al., 1987).

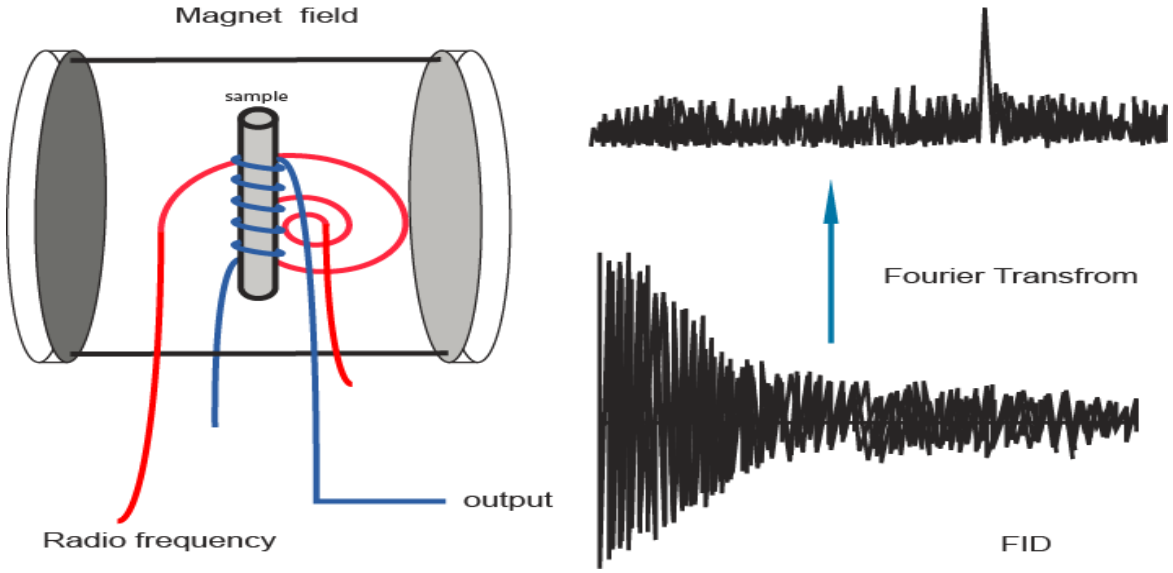

Figure 7. Principle of a Fourier Transform NMR spectrometer: In the Fourier Transform spectrometer, the external magnetic field is kept constant. The sample is placed in the middle of the homogenous magnetic field created by a superconducting coil, and then an electromagnetic radiation with a specific frequency range is applied to tilt the magnetization vector of the sample away from its equilibrium position. The excited magnetization vector precesses around the direction of the magnetic field, which induces an oscillating current in the receiver coil. This signal decreases with time and is referred to as the free induction decay (FID). The FID is a time based signal. In order to get an NMR spectrum the FID is subjected to Fourier transformation to convert it into a function of frequency. This is the way that the NMR spectrum is displayed.

Biomolecular NMR mainly focuses on the nuclei ^{1}H, ^{13}C, ^{15}N and ^{31}P, which are the major components of the biomolecules (Gardner and Kay, 1998). Due to low natural abundance of the ^{13}C and ^{15}N isotopes, it is quite common to use cultivation media enriched in these isotopes to alleviate this problem. The introduction of these NMR active isotopes into biomolecules provides a huge advantage for modern structural NMR (Kanelis et al., 2001; Markley et al., 2003; Torizawa et al., 2004).

1.2.2 Nuclear magnetic spins interaction

1.2.2.1 Chemical shift

When a molecule is put in an external magnetic field, the charged electrons, just like the nuclei, will rotate with a spin producing a magnetic field opposite to the external magnetic field. Thus, the total effective magnetic field that acts on the nuclear spins will be reduced depending on the strength of the local magnetic field. This is called nuclear shielding and gives rise to the phenomenon known as the chemical shift (Brateman, 1986). The chemical shift describes the difference between the resonance frequency of a given nucleus within its local magnetic field relative to the resonance frequency of an arbitrarily chosen reference substance. It explains why NMR can probe the chemical structure of a molecule. The chemical shift for a resonance is normally measured in ppm shifted relative to a reference signal by the **equation 1.3**

$$\delta = \frac{v - v_{ref}}{v_{ref}} \qquad (1.3)$$

v is the resonance frequency of the observed nucleus. v_{ref} is the resonance frequency of the reference substance. Since the difference between the observed resonance frequency and that of a reference substance is usually in hertz, and the frequency of operator in megahertz, δ is expressed in parts per million (ppm) (Jardetsky and Roberts, 1981). Normally the detected frequencies for ^{1}H and ^{13}C nuclei are usually referenced against tetramethylsilane (TMS), which is assigned the chemical shift of zero. Because of the reference, the chemical shift is not dependent of the applied magnetic field anymore which makes it more convenient for comparing (Wishart and Nip, 1998).

The intimate relationship between the chemical shift and the chemical environment of a nucleus provides the basis for the study of chemical and biological molecular structures by NMR. For example, in a protein each individual atom is packed into different chemical environments, dependent on the primary, secondary and tertiary structure. Thus the same type of nucleus from different amino acids or the same amino acid will have different chemical shifts which allows site

specific studies of a molecule by NMR. The difference of protons chemical shifts in a folded protein can be seen in **Figure 8**.

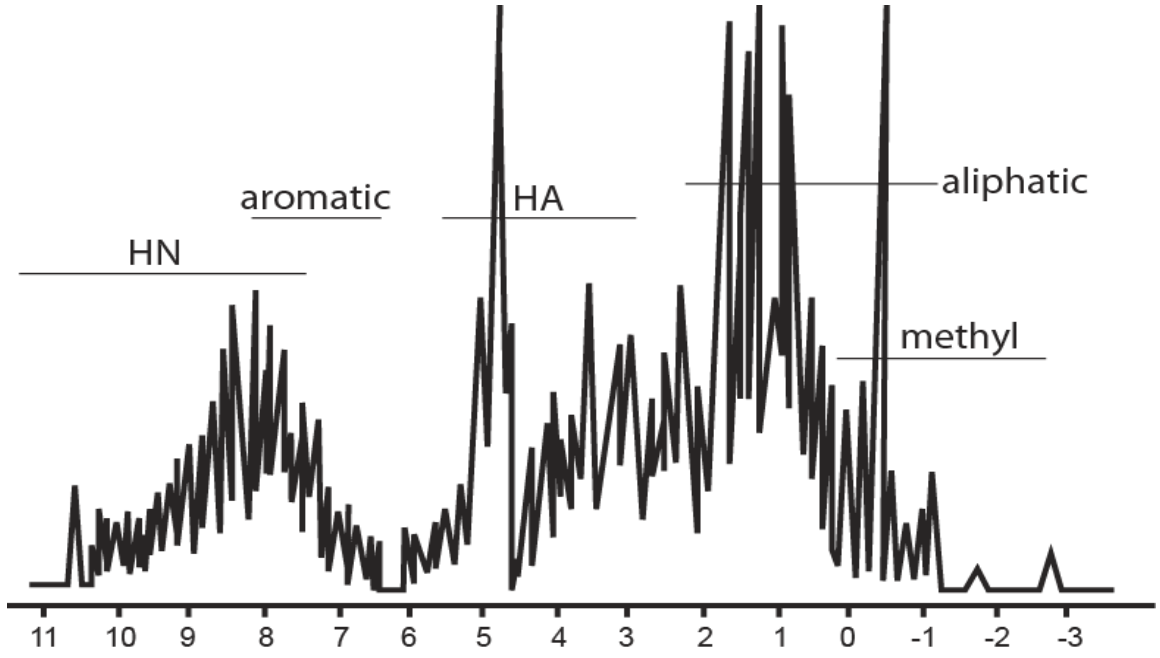

Figure 8. 1H-NMR spectrum of a folded protein: The protons in a folded protein are located in many different chemical environments which leads to a well dispersed spectrum. The chemical shift of amide protons is usually expanded from 6.5 ppm to more than 10 ppm.

The chemical shift can be used to derive secondary structure elements and dihedral φ and ψ, in particular the analysis of the chemical shifts of $^{1}H_{\alpha}$, $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ (Wishart and Sykes, 1994). The plot of the φ vs ψ torsion angles in proteins is called the Ramachandran plot and helps to identify which amino acids are located α-helices or β-strands, respectively. The other way to determine the secondary structure of a protein is the so-called 'secondary chemical shift' (Dalgarno et al., 1983), defined as:

$$\Delta\delta = \delta_{observed} - \delta_{random\ coil} \qquad (1.4)$$

The $\Delta\delta$ values of C_{α} atoms in helices tend to be positive and in β-strands tend to be negative. The $\Delta\delta$ of C_{β} atoms is exactly opposite to that of C_{α} atoms (Wang and Jardetzky, 2002a).

1.2.2.2 Scalar coupling

Scalar coupling, also named spin-spin coupling or J coupling (isotropic part) arises due to electrons participating in bond(s) connecting the nuclei (Cavanagh et al., 2007; Ernst et al., 1987; Levitt, 2008). Two nuclei in a molecule linked by less than three bonds can be seen as coupled in NMR spectra. The magnitude of a scalar coupling is usually larger when nuclei are connected through fewer bonds. From physical point of view, J coupling constant is proportional to the product of gyromagnetic ratios of the involved isotopes. It is not dependent of magnetic field. However it depends on molecular geometry, especially the conformation. In biomolecular NMR, coupling constants are mainly used to determine dihedral angles. Precise measurement of J coupling between H^{α} and H^{N} could be used to determine the φ angle in the peptide bonds based on Karplus function (Evans, 1995; Karplus, 1959, 1963). J coupling is also quite heavily applied in liquid state NMR. It is used for coherence transfer in many experiments such as HSQC (Schleucher et al., 1994)(Bodenhausen and Ruben 1980), COSY (Nagayama, 1981) and TOCSY (Schleucher et al., 1994). For this purpose a delay for J coupling evolution is introduced in the pulse program. In other cases, J coupling also needs to be suppressed. For example, ^{13}C spectra for solid state NMR are usually recorded with ^{1}H decoupling (Ramamoorthy et al., 1999).

1.2.2.3 Dipole-dipole interaction

A nuclear spin in a magnetic field is a magnetic dipole. Dipole-dipole interaction arises due to two or more non-zero spins close to each other in an external field. Any individual nuclear spin works like a small magnet producing its own magnetic field. The effective magnetic field at each spin will be the sum of local field and external field. In a result, all near-by nuclear spins will have an influence on each other. The magnitude of dipolar coupling depends on the inverse third power of the distance between two spins and the product of their gyromagnetic ratios. It can be calculated as:

$$D = -\frac{\mu_0 \gamma_1 \gamma_2 \hbar}{4\pi r^3} (3\cos^2\theta - 1) \quad (1.5)$$

The dipolar coupling is anisotropic in nature. Its magnitude also has orientation dependence through the direction of the vector connecting the two nuclear spins. In solution where the molecules tumble very fast, the dipole-dipole interactions are averaged out and not observed, but it is possible to reintroduce the dipolar coupling in colloidal solutions. This is generally measured as residual dipolar coupling to extract structural information (Cavanagh et al., 2007; Evans, 1995). Another more commonly used method to get structural information is through the nuclear Overhauser effect (NOE), which is the transfer of magnetization between spins coupled by dipole–dipole interactions in a molecule undergoing brownian motion. The intensity of NOE is inversely proportional to the sixth power of the distance between the two nuclear spins. Thus the NOE can be used to estimate the distance between two nuclei (Evans, 1995).

In solid state, molecules are limited in tumbling, resulting in nuclear spins that are heterogeneously maintained in fixed orientation. Therefore the anisotropic dipolar coupling can extremely broaden the line width. On the other hand dipolar couplings are also used in solid state NMR as the primary source for polarization transfer for many measurements (Baldus and Meier, 1996; Li et al., 2012). This will be discussed in the section of solid state NMR.

1.2.3 Solution NMR spectroscopy and structure determination

The phenomenon of magnetic resonance results from the quantum mechanical properties of the nucleus and its interaction with external magnetic field. These quantum mechanical properties depend on the local chemical environment of the nucleus and the magnetic field produced by neighbouring nuclei. The measurement on these properties provides information on orientations, angles and distances in space. Solution NMR spectroscopy of proteins is one of the two main techniques used in structural biology, and it is used both to determine protein structures and to obtain dynamic information. The process of protein structure determination can be divided into five different phases: sample preparation and buffer optimization, spectra measurements, NMR signal assignments, conformational constraints identification and structure calculation (Wider, 2000; Wider and Wüthrich, 1999).

1.2.3.1 Sample preparation and spectra measurements

Typically the proteins studied by solution NMR consist of less than 300 residues. The protein is expressed from bacteria grown in minimal medium enriched in NMR-active isotopes and purified to high purity. Based on 1D ^{1}H-NMR spectra one can easily tell if a protein is folded or aggregated. Different buffer conditions have to be tested to screen the best buffer for measurement. In the next step, a set of heteronuclear multidimensional NMR experiments is recorded. The first experiment is usually a 2D ^{1}H-^{15}N heteronuclear single quantum correlation (HSQC) spectrum (Bodenhausen and Ruben 1980), in which one single is expected for the backbone amide of each amino acid with exception of proline which has no amide-proton. Asparagine and other amino acids could give rise to more than one peak because of their side-chain amide groups. The ^{15}N-HSQC (Bodenhausen and Ruben 1980) is often referred to as the fingerprint of a protein and analysed to decide which direction one should go in later steps (Wider, 2000).

1.2.3.2 Assignments and structure calculation

In order to analyse the NMR data, each resonance must be linked to an atom, or a pair of atoms corresponding to it. This is called assignment including backbone assignment and side chain assignment. Usually the protein is isotopically (^{15}N, ^{13}C) labeled. The distance constraints are extracted from NOESY spectra, in which the intensity of cross peaks has correlation with distance between two nuclei as mentioned in **section 1.2.2.3**. Thus each peak can be converted into distance constraints, normally between 1.8 to 6 angstrom (Wider, 2000). Since intensity-distance relationship is not precise, a distance range is used. Many computer programs have been developed to automatically assign the NOESY peaks such as UNIO (Guerry and Herrmann, 2012), CANDID (Guntert, 2004) and ARIA (Brunger et al., 1998; Rieping et al., 2007). In addition to distance restraints, angle restraints (φ and ψ angles) generated from chemical shifts are used in a restrained molecular dynamics/simulated annealing (MD/SA) structure calculation, which is usually an iterative process. The result is an ensemble of structures, all of which should be consistent with the input NMR data (Bassolino-Klimas et al., 1996).

1.2.4 Relaxation and dynamics in solution

Protein dynamics by NMR mainly refers to time dependent physical or chemical fluctuations in structure involving both inter and intra molecular motions. The timescale ranges from picoseconds to

hours or even longer. Different techniques or methods are used dependent on the timescale. Protein dynamics could affect a range of functions including catalytic activity of enzymes, protein-protein or protein-ligand interactions, as well as biophysical properties such as thermo stability etc. Due to the sensitivity and capability of measuring different scale times, NMR has become the most powerful tool to study dynamics and relaxation. (Kay, 1998, 2005).

1.2.4.1 Relaxation

Relaxation is a term describing the process that nuclear magnetization prepared in a non-equilibrium state tends to come back to the thermal equilibrium distribution. Relaxation combines two different components for isotropic systems in the absence of chemical exchange: Longitudinal relaxation corresponding to longitudinal magnetization recovery and transverse relaxation corresponding to transverse magnetization decay. Longitudinal relaxation, also named spin-lattice relaxation, describes the return of the *z* component of the nuclear spin magnetization back to equilibrium state by release of energy to the surrounding lattice. This process follows an exponential curve. After a time constant T1, the magnetization has recovered to 63% of its equilibrium value. Transverse relaxation, also named spin-spin relaxation, is a complex phenomenon mainly due to a cumulative loss in phase of nuclear spins. It is also described by an exponential curve. After a time constant T2, the transverse magnetization has lost 63 % of its starting value. For large molecules, the line width of an NMR signal can be related to T2 by

$$W_{1/2} = \frac{1}{\pi T_2} \qquad (1.6)$$

As a result, the line width of signals from macromolecules will be much broader than that from small molecules. This phenomenon is the reason why protein structure determination by solution NMR becomes problematic for larger proteins. Broad line width results in low signal intensity and gives rise to heavy signal overlay when many nuclei are present (Cavanagh et al., 2007; Evans, 1995).

1.2.4.2 Chemical exchange

Chemical exchange in general refers to any process that causes a modulation in the isotropic chemical shift. It occurs in the microsecond to millisecond time scale and contains both intramolecular (e.g. conformational change) and intermolecular processes (e.g. binding of ligands to receptor). Studying chemical exchange could provide a lot of biologically meaningful kinetic and interaction parameters (Bain, 2003, 2006). When NMR signals in spectra have been assigned, each signal is a specific reporter for individual nuclei in the protein structure. If there are changes in the environment of a spin in a protein e.g. due to binding of its interaction partner, the chemical shift will change. These changes are significant on the binding site. Therefore, the binding surface on a protein with its interaction partner can be easily identified. In addition, the thermodynamic parameters of such a process can be determined from NMR titration experiments. In this sense, NMR is by far a superior technique compared to other spectroscopy to identify and record these changes for analysis (Culf et al., 2009; Cuperlovic et al., 2000).

1.2.5 Solid-state NMR spectroscopy

Unlike in solution NMR, where the rapid tumbling of molecules in solution leads to averaging of anisotropic interactions, in solid state NMR, in media with no or little mobility (e.g. crystalline, powder, fibril, aggregates) (Jaroniec et al., 2004; McDermott, 2004; Thompson, 2002; Wasmer et al., 2008), anisotropic interactions have a substantial effect on nuclear spins. Two major of these directionally dependent interactions are chemical shift anisotropy (CSA) and dipolar coupling which extremely broaden the line width. However, high-resolution spectra could still be obtained by the introduction of magic angle spinning (MAS) to get the same level of information that could be obtained from solution NMR (Griffin, 1998). Moreover, solid state NMR can be applied without fundamental size limitation. It is widely applied in studying biological systems of protein aggregation, which in their native form are normally insoluble and noncrytalline. Thus solution NMR or X ray can’t be used to obtain the atomic level structure of these systems. The unique capability of solid state NMR makes it a powerful tool to obtain high resolution structural information for many insoluble samples of great biological importance.

1.2.5.1 Magical angle spinning

As discussed in **section 1.2.5**, in solid state NMR the brownian molecular motions are more restricted. Therefore the anisotropic interactions could give rise to a broad line width. However, both dipolar coupling and chemical shift anisotropy contain ($3\cos^2\theta$ -1) term. By spinning the sample at the angle $\theta = 54.74°$ with respect to B_0, the term ($3\cos^2\theta$ -1) equals to 0. This technique is called

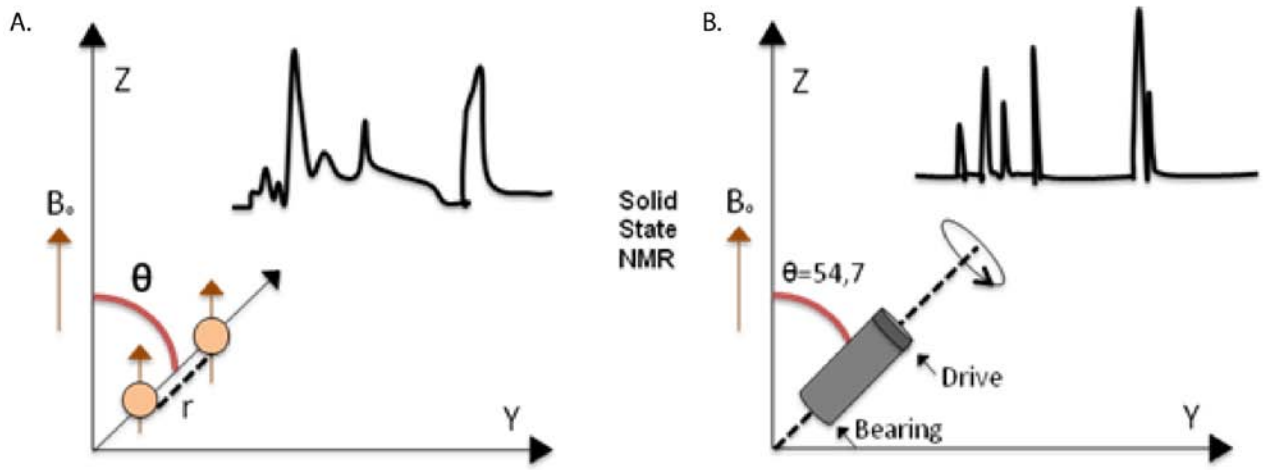

Figure 9. Principle of magic angle spinning in solid state NMR: A. Solid state NMR spectrum without MAS which has very broad peaks due to anisotropic interactions. **B.** Rotation of the sample rotor at magic angle $\theta = 54.74°$ with respect to B_0 leads to high resolution spectrum.

magic angle spinning (MAS). With MAS, the broad line width caused by anisotropic interaction can become narrower, which leads to a significant increase of the resolution of the spectrum (**Figure 9**). However, magic angle spinning (MAS) rate should be larger than or at least equal to the magnitude of the anisotropic interaction to average it to zero. For example, the spinning rate should be $\geq$ 60 kHz to average the ^{1}H-^{1}H homonuclear dipolar coupling. That is the reason that for most solid-state NMR probes designed with spinning rates less than 60 kHz, ^{13}C and ^{15}N signals rather than ^{1}H signal are detected. The speed of the solid state NMR rotor is determined with an optical system. The magic angle can be calibrated using a sample of KBr (Morcombe and Zilm, 2003; Saitô et al., 2006; Thompson, 2002).

1.2.5.2 Cross polarization

Solid state NMR typically detects ^{13}C and ^{15}N nuclei, which have an inherently low sensitivity due to their lower gyromagnetic ratios. In contrast, protons are more sensitive NMR nuclei because of high gyromagnetic ratio and high natural abundance. In solid state NMR, the weak C/N signal could be enhanced by cross polarization (CP) in addition to isotopic labeling of the sample (Pines et al., 1972). Cross polarization is a technique that transfers magnetization from a sensitive spin I like ^{1}H to less sensitive spins S such as ^{13}C and ^{15}N to enhance the S/N ratio potentially by a factor of γ_I/γ_S (Pines et al., 1973). Since abundant spins are strongly dipolarly coupled, they have rapid spin-lattice relaxation, which is decisive for the duration of the recycle delay in an NMR experiment. As a result,

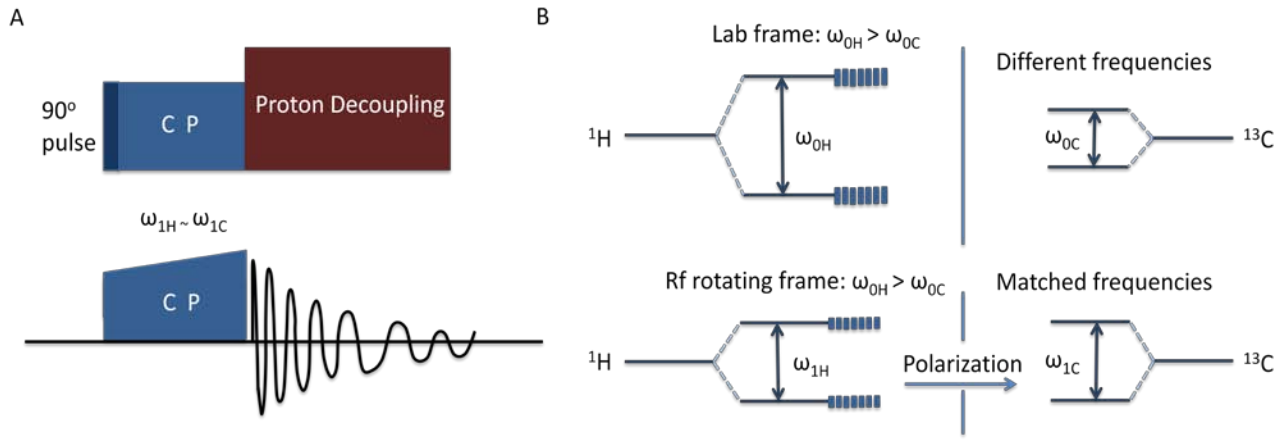

Figure 10. Pulse sequence for cross polarization (CP) experiment in MAS solid state NMR: A. Black bar is a 90° proton excitation pulse, which can be transferred to observed nuclei (C/N) if the proton and observed nuclei magnetization precessed at the same rate. **B.** To obtain efficient cross polarization, the Hartmann-Hahn match should be set properly. The radio frequency field of observed nuclei should be set equal to that of proton by adjust the power on each channels based on equation 1.7 to allow for polarization transfer (Saitô et al., 2006).

cross polarization can also save a lot of measuring time. The standard pulse sequence for cross polarization is shown in **Figure 10**. The abundant nucleus is first excited by a 90° pulse, and then its energy is transferred to the observed nucleus by using a simultaneous low power pulse on both the abundant nucleus and the observed nucleus channels. The key to have efficient cross polarization is to set the Hartmann-Hahn match properly. The radio frequencies on both channels have to be matched according to this equation:

$$\gamma_{13C} B_{13C} - \gamma_{1H} B_{1H} = n \omega_r \quad (1.7)$$

If the cross polarization is done without magic angle spinning, $n \omega_r$ is 0. n is $0, \pm 1, \pm 2, \ldots$ ω_r is the magic angle spinning frequency (Saitô et al., 2006).

1.2.5.3 Recoupling in MAS solid state NMR

Magic angle spinning narrows the spectral line width leading to high resolution and sensitivity. It provides the basic requirement for spectra analysis. However, a lot of information regarding to inter-nuclear distances contained in anisotropic interactions is also averaged out. In order to selectively retrieve this part of information in the presence of MAS, many recoupling techniques have been developed (Dusold and Sebald, 2000; Griffin, 1998; Hohwy et al., 2002; Hohwy et al., 1999). Among these techniques, the anisotropic interaction is usually recovered by applying RF pulse trains. Two of the most illustrative examples are rotational echo double resonance (REDOR) and transferred echo double resonance (TEDOR) (van Eck and Veeman, 1993) . The recovery of dipolar couplings can be easily incorporated into 2D spectra to generate cross peaks between two spatially close nuclear spins. Thus the combination of MAS and recoupling has been widely used in many experiments like dipolar assisted rotational resonance (DARR) spectra to get distance constraints and other structural information (Takegoshi et al., 2001).

1.2.6 Protein structure determination by solid state NMR

MAS solid state NMR experiments are well suited to obtain atomic level resolution structures of proteins that cannot be studied by X-ray crystallography or solution NMR, which includes fibrillar proteins such as amyloid fibrils or some membrane proteins. Since the first high-resolution protein structure was solved by solid state NMR in 2002 (Castellani et al., 2002), a lot of progress has been made in structure determination of small proteins. High resolution spectra of the prion domain of HET-s were reported in 2005 (Ritter et al., 2005), and its atomic resolution structure was published in 2008 (Wasmer et al., 2008). Sparse labeling strategies using either 1,3-^{13}C and 2-^{13}C glycerol or 1-

^{13}C glucose and 2-^{13}C glucose yielded higher resolution NMR spectra and new assignment methods based on the specific labeling pattern (Castellani et al., 2002; Loquet et al., 2011; Lundstrom et al., 2007). Novel NMR pulse techniques for acquiring structural information and new software tools are also constantly being developed allowing today the complete assignment of NMR signals to their corresponding atoms, which is prerequisite of structure calculation. With this fast progress, solid state NMR is becoming a more and more valuable biophysical tool for studying many biological systems that were considered difficult before.

1.2.6.1 Labeling strategy

^{1}H detection in solid state NMR is still a challenge for most MAS NMR probes, due to the design of the speed for magic angle spinning. In many cases, ^{13}C and ^{15}N signals are directly detected even though both of them have lower sensitivity. Thus, solid state NMR spectroscopy of proteins heavily depends on labeling with the stable isotopes ^{13}C or ^{15}N. A variety of labeling strategies have been developed for different requirements in practice (Verardi et al., 2012).

For peptides less than 40 amino acids, solid phase synthesis can be applied (Verardi et al., 2012). Therefore, it is very convenient to introduce specifically isotope labeled amino acids at any necessary position. However, for large proteins, the most commonly used way is expressing them in *E.coli* or yeast. In these systems, the labeling pattern critically depends on the sole nitrogen and carbon source in the medium. There are three main approaches for protein labeling: selective labeling, sparse labeling and uniform labeling. All of the three ways are used extensively in solid state NMR spectroscopy for protein research. Selective labeling refers to incorporation of a single or several types of labeled amino acid in the medium or reversely wiping out specific types of amino acids by adding in unlabeled amino acid in the enriched isotope medium. Sparse labeling refers to using precursors with a specific isotope labeling pattern such as 1- or 2- ^{13} C glucose (Lundstrom et al., 2007), 1,3- or 2- ^{13}C glycerol (Hong and Jakes, 1999) to specifically label certain carbons within each amino acid type. Previous work demonstrates that sparse labeling strategies could give rise to high resolution spectra and farther spin diffusion-based transfer of magnetization which provides longer distance constraints (Castellani et al., 2002). Mixing of monomeric molecules with different sparse labeling patterns to form the final complex offers the possibility to distinguish inter- and intra-molecular distance restraints, which is extremely critical for determining the quaternary structure of homotypic protein assemblies. The uniform labeling strategy is the simplest and most efficient way

by using the uniformly ^{13}C labeled glucose and ^{15}N labeled ammonium chloride as the sole C or N source in the growth medium. In some cases, one uniform labeled protein is sufficient for structure determination (Shahid et al., 2012a). However, it is only applicable for determining the structure of a protein with relatively smaller molecule weight, due to the ambiguity problem causing by broad line width.

1.2.6.2 3D structure determination by solid state NMR

As in solution NMR, for a further analysis of information encoded in NMR spectra, nearly complete backbone and side chain assignments are prerequisite. In solid state NMR, the protein is normally ^{13}C and ^{15}N labeled so that the entire protein is almost a continuous system. Especially the backbone is one long series of coupled nuclei. Three dimensional NCACX (Pauli et al., 2001), NCOCX (Pauli et al., 2001) and CONCA (Li et al., 2007) experiments have been established to detect these continuously coupled nuclei by linking ^{13}C resonances to the backbone ^{15}N resonances providing a way for sequential assignment. 2D carbon-carbon correlation spectra could be analysed to find out the fingerprint patterns of various types of amino acids. By combination of the two- and three- dimensional experiments, nearly complete assignment of carbon and nitrogen could be achieved (Shahid et al., 2012b; Shi et al., 2011).

For determination of a protein structure, medium and long range distance constraints are required. 2D and 3D proton drive spin diffusion (PDSD) experiments (Szeverenyi, N. M. et al. 1982) with different mixing time could be explored to extract large numbers of distance constraints. However, in solid state NMR, the polarization transfer is dominated by strong dipolar couplings which arise mainly from covalently linked carbons, while the through-space couplings which produce the useful constraints are heavily attenuated. To alleviate this problem, sparse labeling strategies could be used to label only one of two adjacent carbons. Consequently, this method permits longer magnetization transfer to access long range constraints.

To solve a tertiary and quaternary structure of a protein, it is important to distinguish inter- and intra-molecular constraints (Loquet et al., 2012). Intra-molecular constraints contain information about how the individual protomer is folded. Intermolecular constraints contain information on how the protomers are assembled into fibrils or a complex. The most common current way is use of mixed isotopically labeled samples such as diluted samples, which is a mixture of unlabeled sample with uniformly labeled sample and mixed sample which is mixture of monomers with different sparse

labeled pattern to form quaternary structure. Often, atomic resolution NMR data is combined with lower resolution structural information obtained by other techniques like electron microscopy for structure calculation (Wright, 2010). Till now, several computer programs such as ARIA (Rieping et al., 2007), XPLOR-NIH (Schwieters et al., 2003) are available to handle C-C or N-C distance constraints together with dihedral angle extracted from chemical shift data for tertiary structure calculation. The quaternary structure calculation is a challenging task due to various, often unknown symmetries of the protein complexes and lacking of matured software tools to handle this symmetry problem.

1.3 Contribution from other techniques to protein structure determination

As more and more protein structures have been determined by NMR spectroscopy and X-ray crystallography, the need to understand how single molecules form a function complex grows. Solid state NMR becomes more and more powerful in this aspect. However, electron microscopy and other techniques such as X-ray fiber diffraction and small angle X-ray scattering (SAXS) could also provide complementary structural information, allowing integration of structural information from different levels for comprehensive description of protein function in molecular detail (Blanchet and Svergun, 2013).

1.3.1 Electron microscopy

The principles of transmission electron microscopy (TEM) are similar to light microscopy. The source for illumination is a cathode which emits electrons rather than light. Electrons are then accelerated by an anode to form an electron beam line travelling through a vacuum down to the specimen. Scattered or unscattered electrons by the specimen are then collected and focused by the magnetic lenses for camera detection. The structural information of the specimen is encoded in the scattered electrons. TEM images are 2D projections of the specimen and affected by contrast transfer function (CTF) of the TEM due to aberrations of the lenses and defocusing used in imaging. Thus the estimation of and correction for CTF are necessary to get a faithfully representative projection image (Voortman et al., 2012).

1.3.1.1 Structural information available from electron microscopy

Measurements of fibril diameters could be easily obtained by transmission electron microscopy which can be used for structure determination. The handedness of helical fibrils could be measured in ideal cases by titled beam-TEM or metal shadowing. Symmetry information of helical assemblies could be extracted from electron diffraction patterns, which is also named power spectra. Based on the Bessel order (n), layer line number (l) and the helix selection rule, the helix indexing could be achieved. Relatively low resolution 3D structures of viral capsids (Clare and Orlova, 2010; Zhang et al., 2013) and fibrils (Al-Khayat et al., 2008; Mun et al., 2011) could be reconstructed via TEM or – with higher resolution – from cryo-EM images. These types of information are difficult to obtain from NMR spectroscopy, which only provides spatially highly localized, high-resolution information, and are thus highly valuable for atomic level structure calculations of supramolecules by NMR.

1.3.1.2 3D reconstruction of helical assemblies

Many proteins exist as helical polymers. These include microtubules, actins, viral capsids and so on. The abundance of helices might be a consequence of their simple but stable assembly from asymmetric units in nature. Mathematically, a helix is described in Cartesian coordinates by three equations $X = r \cos \phi$, $Y = r \sin \phi$, and $Z = z$ (**Figure 11**). It looks like a circle gradually rotating and rising along the Z axis. For each subunit, the rotation is $\Delta\Phi$ and the rise is Δz. Generally, the helix will repeat at a certain distance c along Z axis, as described by exactly u subunits in t turns. These values can be calculated from the following two equations:

$$\Delta z = c/u \tag{1.8}$$

$$\Delta\Phi = 2\pi\, t/u \tag{1.9}$$

The helical symmetry is relatively simple, because it can be converted into a 2D lattice which is helpful to understand the process of 3D reconstruction (**Figure 11**). If a sheet with the 2D lattice is rolled into a cylinder, the lines on the sheet will spiral along the axis of the cylinder at constant radius r. Two parameters are needed to describe the 2D lattice: the repeat distance between two

neighboured lattice lines along the helical axis (l) and the number of lattice lines that pass through the equator (the order n). For a certain 2D lattice, it can be described by different sets of l and n,

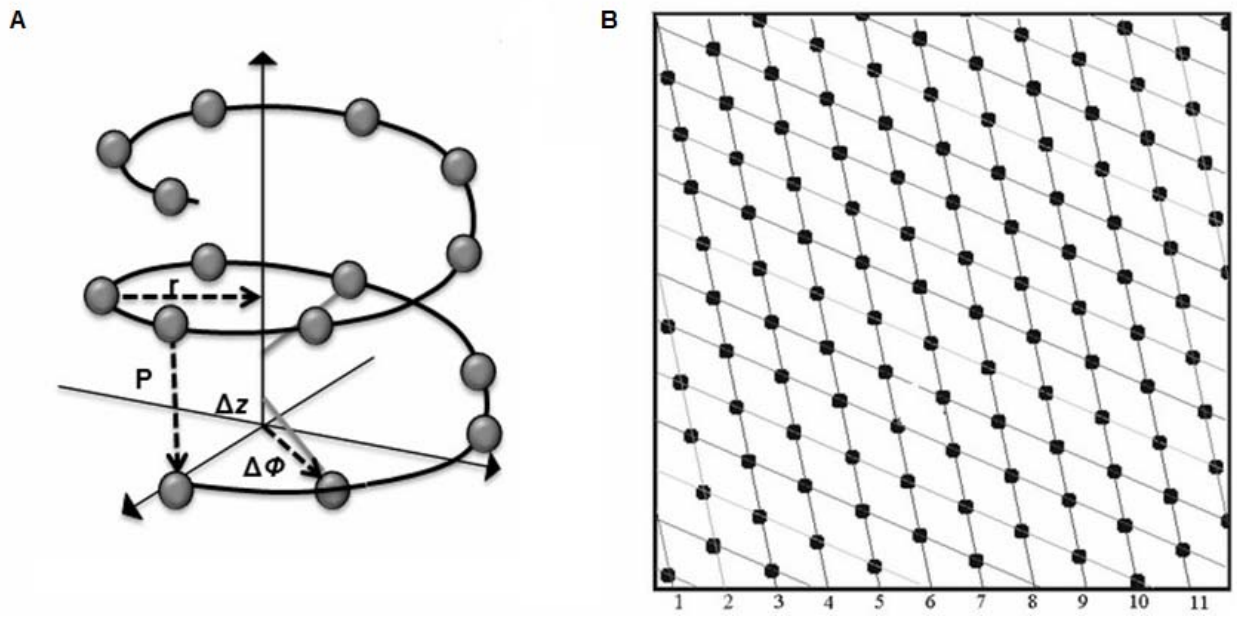

Figure 11. Mathematic description of helix and its related 2 D lattice A: the geometry of a helix. *P* is the pitch of the helix which is the height of one complete helix turn along the axis. *r* is the radius of the helix. $\Delta\Phi$ and Δz are the incremental rotation and translation between each subunit. **B.** A helical structure can be thought of as a 2D lattice rolled into a tube.

however for any set of parameters they are restricted by a helix selection rule:

$$L = um + tn \qquad (1.10)$$

where m can be any integer. The helix selection rule is a linear equation with integral coefficients and two variables m and n. It is the basis for helical polymer reconstruction (Meng et al., 2011; Zhang et al., 2013).

The first 3D structure reconstructed from 2D images was a helical phage tail (De Rosier and Klug, 1968), because in helices a single subunit is repeated many time giving projections from different angles. The symmetry itself also rises restriction for the final 3D density map. These properties make

the reconstruction much easier. Till now, several methods have been developed for helical reconstruction. One of the most often used today is the iterative helical real space reconstruction (IHRSR) method (Egelman, 2007), which is an algorithm for single particle reconstruction of helical polymers. It starts with a measured or a reasonable guessed symmetry, performing iterative steps to refine a featureless solid cylinder into the final real helical structure of the polymers (Diaz et al., 2010; Egelman, 2010).

1.4 Aim of the project

RLR family members RIG-I and MDA5 represent a conserved family of vertebrate innate immune receptors which could recognize viral RNAs in the early stage of infection, initiating the signaling cascade by interacting with the mitochondrial antiviral-signaling (MAVS) protein. Recent work revealed that upon activation by 2CARDs of RIG-I or MDA5, MAVS proteins aggregate on the surface of mitochondria and form protein filaments. The formation of MAVS filaments results in a gain of function, and this higher-order assemblies could then draw more inactive MAVS proteins into active aggregates, leading to the formation of more filaments and amplifying the signal cascades. This process is considered to be a new paradigm of signal transduction (Wu, 2013). However, the underlying structural basis of MAVS activation has yet to be described.

The first objective of this thesis is to determine the structure and dynamics of the monomeric, inactive form of MAVS CARD by solution state NMR, in order to gain novel insights into its mechanism of activation, and to identify the interaction site of its binding partner RIG-I CARD.

The second objective is to carry out a detailed study on the structure of active fibrillar MAVS CARD by solid state NMR, in order to determine what kind of conformationals changes might occur during MAVS assembly, and to identify the chemical interactions between individual protomers within the filament at atomic resolution. These objectives are expected to help unravel the mechanism of MAVS activation, including the role RIG-I CARD domains might play as a 'seed' for MAVS-CARD aggregation. Atomic resolution knowledge of homo- and heterotypic CARD-CARD interaction surfaces will also provide the basis for understanding how viral proteins interfere with MAVS signal transduction.

2 Materials and Methods

2.1 Material

The chemicals used in this project were of quality p.a. (pro analysis). All chemicals were purchased from the following companies: Eurofins MWG Operon, Fermentas, GE-Healthcare, Invitrogen, Merck, Millipore, Qiagen, Roche, Sigma-Aldrich and Omnilab, Fluka, Stratagene.

2.1.1 Enzymes, Kits and Molecular weight standards

Enzymes, Standard Markers, and Kits were used as describe by the corresponding company.

Table 2. Enzymes used in this project.

Enzyme	Company
BamHI, EcoRI, DpnI, NdeI, HindIII	Fermentas
DNAse	Merck
Platinum Pfx Polymerase	Invitrogen
Pfu Turbo Polymerase	Agilent
T4 DNA Ligase	New England Biolabs
Hen white-egg lysozyme	Fluka

Table 3. Standard marker used in this project.

Name	Company
Smart Ladder	Eurogentec
PageRuler^{TM} Prestained Protein ladder	Fermentas
PageRuler^{TM} unstained Protein ladder	Fermentas

Table 4. Kits used in this project.

Name	Company
QIAquick Gel Extraction Kit	Qiagen
QIAquick PCR Purification Kit	Qiagen
QIAprep Spin Miniprep Kit	Qiagen
QIAprep Spin Midiprep Kit	Qiagen
MitoTracker® Red CMXRos	Invitrogen

2.1.2 Plasmids

The recombinant plasmids used in this project were listed in the following table (**Table 5**). Plasmids were constructed from pET-21a and pcDNA3.1a. Both of them carried ampicillin resistance gene. MAVS and RIG-I genes were synthesized from the company GENEART, where the codons optimization was done for improving expression in *E.coli*.

Table 5. Plasmids used in this project.

Plasmid	Source
T-vector-MAVS$_{1\text{-}540}$	GENEART
pETRO2.11 SUMO*-MAVS$_{1\text{-}100}$	Cloned from T-Vector MAVS $_{1\text{-}540}$
pETRO2.11 SUMO*-MAVS$_{1\text{-}152}$	Cloned from T-Vector MAVS $_{1\text{-}540}$
pETRO2.11 SUMO*-MAVS$_{1\text{-}175}$	Cloned from T-Vector MAVS $_{1\text{-}540}$
pETRO2.11 SUMO*-MAVS$_{1\text{-}100\ K7S\ K10S\ R14S}$	Cloned from T-Vector MAVS $_{1\text{-}540}$
pETRO2.11 SUMO*-MAVS$_{1\text{-}100\ Y9E}$	Cloned from T-Vector MAVS $_{1\text{-}540}$
pETRO2.11 SUMO*-MAVS$_{1\text{-}100\ D23S\ E28S}$	Cloned from T-Vector MAVS $_{1\text{-}540}$
pETRO2.11 SUMO*-MAVS$_{1\text{-}100\ N21A\ D23A}$	Cloned from T-Vector MAVS $_{1\text{-}540}$

pETRO2.11 SUMO*-MAVS$_{1\text{-}100\ R37S\ R41S}$	Cloned from T-Vector MAVS $_{1\text{-}540}$
pETRO2.11 SUMO*-MAVS$_{1\text{-}100\ R64S\ R65S}$	Cloned from T-Vector MAVS $_{1\text{-}540}$
pETRO2.11 SUMO*-MAVS$_{1\text{-}100\ E80S\ E83S}$	Cloned from T-Vector MAVS $_{1\text{-}540}$
pETRO2.11 SUMO*-MAVS$_{1\text{-}100\ D86S\ E87S}$	Cloned from T-Vector MAVS $_{1\text{-}540}$
pETRO2.11 SUMO*-MAVS$_{1\text{-}100\ D23S}$	Cloned from T-Vector MAVS $_{1\text{-}540}$
pETRO2.11 SUMO*-MAVS$_{1\text{-}100\ E28S}$	Cloned from T-Vector MAVS $_{1\text{-}540}$
pETRO2.11 SUMO*-MAVS$_{1\text{-}100\ R37S}$	Cloned from T-Vector MAVS $_{1\text{-}540}$
pETRO2.11 SUMO*-MAVS$_{1\text{-}100\ R41S}$	Cloned from T-Vector MAVS $_{1\text{-}540}$
pETRO2.11 SUMO*-MAVS$_{1\text{-}100\ R64S}$	Cloned from T-Vector MAVS $_{1\text{-}540}$
pETRO2.11 SUMO*-MAVS$_{1\text{-}100\ R65S}$	Cloned from T-Vector MAVS $_{1\text{-}540}$
pETRO2.11 SUMO*-MAVS$_{1\text{-}100\ W56A}$	Cloned from T-Vector MAVS $_{1\text{-}540}$
pETRO2.11 SUMO*-MAVS$_{1\text{-}100\ V25S}$	Cloned from T-Vector MAVS $_{1\text{-}540}$
pETRO2.11 SUMO*-MAVS$_{1\text{-}100\ S49A}$	Cloned from T-Vector MAVS $_{1\text{-}540}$
pcDNA3.1 HA-MAVS$_{1\text{-}540}$	Renate König Paul-Ehrlich-Institut
pcDNA3.1 HA-MAVS$_{1\text{-}540\ Y9E}$	Cloned from Renate König's Vector
pcDNA3.1 HA-MAVS$_{1\text{-}540\ D23S\ E26S}$	Cloned from Renate König's Vector
pcDNA3.1 HA-MAVS$_{1\text{-}540\ N21A\ D23A}$	Cloned from Renate König's Vector
pcDNA3.1 HA-MAVS$_{1\text{-}540\ R37S\ R41S}$	Cloned from Renate König's Vector
pcDNA3.1 HA-MAVS$_{1\text{-}540\ R64S\ R65S}$	Cloned from Renate König's Vector
pcDNA3.1 HA-MAVS$_{1\text{-}540\ E80A\ E83A}$	Cloned from Renate König's Vector
pcDNA3.1 HA-MAVS$_{1\text{-}540\ V25S}$	Cloned from Renate König's Vector
pcDNA3.1 HA-MAVS$_{1\text{-}540\ W56A}$	Cloned from Renate König's Vector
pcDNA3.1 MiniSOG	Roger Tsien UCSD
pcDNA3.1 MiniSOG-MAVS$_{1\text{-}540}$	Cloned from Renate König's Vector

T-vector-RIG-I $_{1-191}$	GENEART
pETRO2.11 SUMO*-RIG-I $_{1-172}$	Cloned from T-vector-RIG-I $_{1-191}$

2.1.3 Oligonucleotides

Oligonucleotides (**Table 6**) were used as primers for cloning and site-directed mutagenesis. They were all purchased from MWG. For mutagenesis, only forward primers were listed in this table.

Table 6. Oligonucleotides used in this project.

Cloning	**Sequence**
RIG-I $_{1-172}$ Forward Primer	cg *ggatcc* atg acc acc gaa cag
RIG-I $_{1-172}$ Reverse Primer	ccg *gaattc* tta acc cag ggt ttt cgg cca
MAVS$_{1-152}$ Forward Primer	cg *ggatcc* atg ccg ttt gct gaa g
MAVS$_{1-152}$ Reverse Primer	ccg *aagctt* tta gga ctc tgg cgc ctg ggt
MAVS$_{1-175}$ Reverse Primer	ccg *aagctt* tta gcc acc atc tgg att cct
MiniSOG Forward Primer	cta *gctagc* atg gag aaa agt ttc g
MiniSOG Reverse Primer	ccc *aagctt* tcc atc cag ctg cac
HA-MAVS$_{1-540}$ Forward Primer	ccc *aagctt* atg tac cca tac gat gtt cca gat tac gct atg ccg ttt gct gaa gac
HA-MAVS$_{1-540}$ Reverse Primer	tgc *tctaga* cta gtg cag acg ccg ccg gt
Mutagenesis	**Forward primers**
MAVS$_{1-100\ K7S\ K10S\ R14S}$	t gca gaa gat tca acc tat tca tac atc tgc tcc aac ttc tcc a
MAVS$_{1-100\ Y9E}$	gca gaa gat aaa acc gaa aaa tac atc tgc cgc
MAVS$_{1-100\ D23S\ E25S}$	c aat ttc tgc aat gtt tct gtt gtg tca att ctg ccg tat ctg c

$MAVS_{1\text{-}100\ N21A\ D23A}$	c tcc aat ttc tgc gcg gtt gcg gtt gtg gaa att ctg ccg tat ctg
$MAVS_{1\text{-}100\ R37S\ R41S}$	cg tgt ctg acc gca tct gat cag gat tct ctg cgt gca acc tg
$MAVS_{1\text{-}100\ R64S\ R65S}$	t aat acc ctg cag tct tct ccg ggt tgg gtt gaa
$MAVS_{1\text{-}100\ E80S\ E83S}$	gca ctg cgt ggt tgt tca ctg gtt tct ctg gca gat gaa gtt g
$MAVS_{1\text{-}100\ D86S\ E87S}$	ctg gtt gat ctg gca tct tca gtt gca agc gtt tat c
$MAVS_{1\text{-}100\ D23S}$	c aat ttc tgc aat gtt tct gtt gtg gaa att ctg ccg tat ctg c
$MAVS_{1\text{-}100\ E28S}$	c aat ttc tgc aat gtt gat gtt gtg tca att ctg ccg tat ctg c
$MAVS_{1\text{-}100\ R37S}$	cg tgt ctg acc gca tct gat cag gat cgt ctg cgt gca acc tg
$MAVS_{1\text{-}100\ R41S}$	cg tgt ctg acc gca cgt gat cag gat tct ctg cgt gca acc tg
$MAVS_{1\text{-}100\ R64S}$	t aat acc ctg cag tct cgt ccg ggt tgg gtt gaa
$MAVS_{1\text{-}100\ R65S}$	t aat acc ctg cag cgt tct ccg ggt tgg gtt gaa
$MAVS_{1\text{-}100\ W56A}$	cgt gat acc ctg gcg cac ctg ttt aat
$MAVS_{1\text{-}100\ V25S}$	c aat ttc tgc aat gtt gat gtt tcg gaa att ctg ccg tat ctg c
$MAVS_{1\text{-}100\ S49A}$	ca acc tgt acc ctg gca ggt aat cgt gat ac
$MAVS_{1\text{-}540\ Y9E}$	gct gaa gac aag acc gaa aag tat atc tgc cgc
$MAVS_{1\text{-}540\ D23S\ E26S}$	c aat ttt tgc aat gtg tct gtt gta tca att ctg cct tac ctg c
$MAVS_{1\text{-}540\ N21A\ D23A}$	c agc aat ttt tgc gcg gtg gcg gtt gta gag att ctg cct tac
$MAVS_{1\text{-}540\ R37S\ R41S}$	cc tgc ctc aca gca tct gac cag gat tct ctg cgg gcc acc tg
$MAVS_{1\text{-}540\ R64S\ R65S}$	c aat acc ctt cag tct tct ccc ggc tgg gtg gag
$MAVS_{1\text{-}540\ E80A\ E83A}$	gca ctg agg ggc tgt gca cta gtt gca ctc gcg gac gaa gtg g
$MAVS_{1\text{-}540\ V25S}$	c aat ttt tgc aat gtg gat gtt tcg gag att ctg cct tac ctg c
$MAVS_{1\text{-}540\ W56A}$	cgg gac acc ctc gcg cat ctc ttc aat acc ctt ca

2.1.4 Bacterial strains

DH5α strain was used in cloning work for producing and amplifying the plasmid, and T7 strain was used for protein expression. The genotypes of these two strains were listed in **Table 7**.

Table 7. Bacterial strains used in this project.

Strains	Genotype
T7	*fhuA2 lacZ::T7 gene1 [lon] ompT gal sulA11 R(mcr-73:: miniTn10--Tet^S) 2 [dcm] R(zgb-210::Tn10--Tet^S) endA1 Δ(mcrC-mrr) 114::IS10*
DH5α	*F– endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG Φ80dlacZΔM15 Δ(lacZYA-argF)U169, hsdR17(rK⁻mK⁺), λ–*

2.1.5 Media and Buffer

The media and buffer that used in this project are summarized in **Table 8**. The media were sterilized by autoclaving at 121 °C, 2 bar for 20 min. The buffer were filtered through membrane with pore width 0.2 µM. Glucose solution for minimal media were also filtered in the same way as other buffer.

Table 8. Media and Buffer used in this project.

Media	Composition
Lysogeny Broth (LB)	1% tryptone, 1% NaCl, 0.5% yeast extract, pH 7.4
Terrific Broth (TB)	1.2% tryptone, 2.4% yeast extract, 0.4% glycerol, 17mM KH_2PO_4, 72 mM K_2HPO_4
CN040	50mM Na_2HPO_4, 50mM KH_2PO_4, 1g/L NH_4Cl, 4g/L Glucose, 5mM Na_2SO_4, 2mM $MgSO_4$, 1% 100xMEM Vitamins (Sigma-Aldrich) 0,02 % 5000x trace metal solution

LB-Agar	LB media adding with1.5 % agar
Buffer	**Composition**
Trace metal solution	100µM $FeCl_3$, 40µM $CaCl_2$, 20µM $MnSO_4$, 20µM $ZnSO_4$, 4µM $CoCl_2$, 4µM $CuCl_2$, 4µM $NiCl_2$, 4µM Na_2MoO_4, 4µM H_3BO_3
Lysis buffer	20mM TRIS-HCl, 50mM NaCl, 1mM DTT, pH 8
Wash buffer	20mM TRIS-HCl, 50mM NaCl, 1mM DTT, 30mM Imidazole, pH 8
Elution buffer	20mM TRIS-HCl, 50mM NaCl, 1mM DTT, 500mM Imidazole, pH 8
6 ×DNA loading buffer	10mM TRIS/HCl pH 7.4, 25mM EDTA, 30% glycerol, 0.4 % OrangeG
4 ×SDS-PAGE loading buffer	8% SDS, 0,002% Bromphenolblau, 40% Glycerin, 20%β-Mercaptoethanol, 0.25M TRIS-HCl pH 6.8
20×SDS-PAGE running buffer	1M MES, 1M TRIS-Base, 70mM SDS, 20mM EDTA
SDS-PAGE staining solution	Instant blue
Buffer A (pH 3)	50mM Na_2HPO_4/ NaH_2PO_4 , 1mM DTT, pH 3
Buffer B (pH 7)	50mM Na_2HPO_4/ NaH_2PO_4 , 1mM DTT, pH 7
EM staining solution	Aqueous 0.5%-4% uranyl acetate solution, pH 4.5 Aqueous 0.5%-3% phosphotungstic acid solution Aqueous 1%-10% ammonium molybdate
EM washing solution	TE-buffer (20 mM TRIS, 1 mM EDTA, pH 6.9)

2.2 Methods

2.2.1 Cloning and mutagenesis

The sequences of human MAVS CARD (1-100) and RIG-I (1-172) were obtained from NCBI database. The coding regions were optimized and synthesized by Geneart / Life Technologies Inc. Both genes were cloned into modified pET21-d vector with N-terminal six-histidine tag and SUMO* upstream of the initiation codon of MAVS/RIG-I gene. The fused protein therefore included N-terminally His-tagged SUMO* and target proteins. Mutants of MAVS were made based on the standard protocol for site directed mutagenesis using *Pfu* Turbo Polymerase. Full length MAVS used for in-vivo studies were cloned under HA tag and miniSOG tag with its native sequence. The restriction enzyme sites used for cloning were indicated in the table of primers (**Table 6**). All of the recombinant constructs were confirmed by restriction enzyme digestion and sequencing in the company Eurofins MWG Operon.

2.2.2 Protein production and purification

The recombinant plasmids were transformed into T7 Express Competent E. coli for protein expression. As in typical experiments, uniformly ^{15}N /^{13}C labeled protein were prepared by growing cells in CN040 minimal medium containing 1 mg/ml $^{15}NH_4Cl$ and 4 mg/ml $^{13}C6$- glucose at 37 °C. When the OD600 of the cell culture reached 1.0, the cells were transferred to a 30 °C incubator and protein expression was induced by using a final concentration of 0.5 mM of Isopropyl β-D-1-thiogalactopyranoside (IPTG). After additional 5 h of incubation, cells were harvested by centrifugation at 5000 g for 20 min. The pellet was then resuspended in 50 ml lysis buffer with around 0.1 mg/ml lysozyme and 0.01 mg/ml DNase. Efficient cell lysis was performed by using homoginizer for two times at 20 Kpsi at 4 °C. The soluble bacterial lysates were separated from cell debris and other components by centrifugation at 16000 g for 45 min. After centrifugation, targeted protein was purified via Ni-NTA affinity chromatography. The proteins were eluted down using elution buffer containing 500 mM imidazole and the elution fraction were analyzed on the SDS gel.

For wild type SUMO*-MAVS-$CARD_{1-100}$, we used a different method for purification. After the separation of the supernatant with cell debris and other insoluble components, ultracentrifugation was then applied for the supernatant at 100,000 g for 2 h. Self-assembled MAVS CARD was collected in the pellet. This pellet was dissolved again in 20 mM TRIS, 50 mM NaCl, 2 mM DTT pH 8 buffer and simultaneously cleaved with SUMO* protease at 4 °C overnight. The sample was

loaded onto the Ni-NTA column (Qiagen). His-tagged SUMO* bound to the column while MAVS CARD flowed through were further purified with a Superdex 200 column with 20 mM TRIS, 50 mM

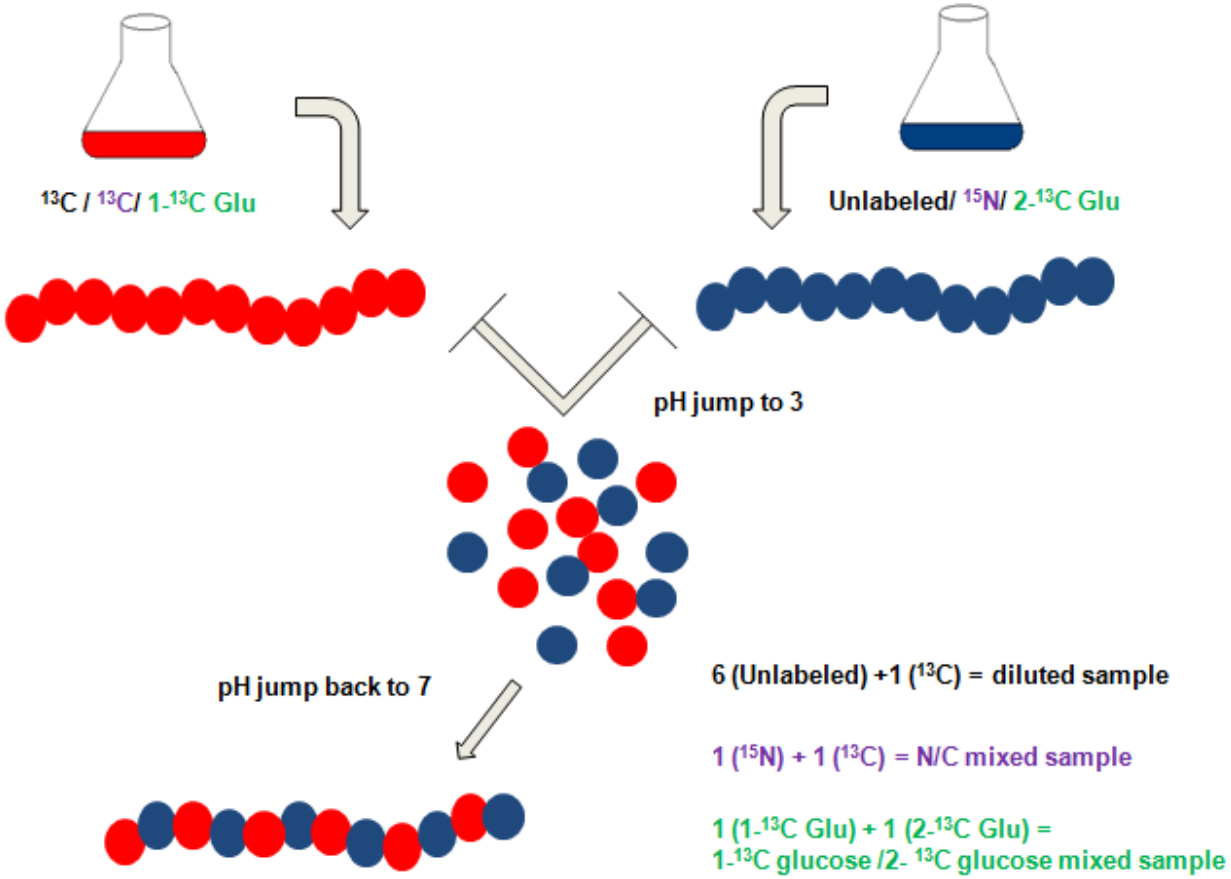

Figure 12. Scheme for preparation of diluted, $^{15}N/^{13}C$ and 1-^{13}C glucose /2- ^{13}C glucose labeled mixed sample. MAVS-CARD$_{1-100}$ was first expressed in the media with different isotopes. After the purification of differently labeled MAVS-CARDs$_{1-100}$, the pH of the buffer was changed to 3 to make the proteins into monomers. Then differently labeled MAVS-CARDs$_{1-100}$ were mixed and the pH of the buffer was brought back to 7 to make diluted, $^{15}N/^{13}C$ and 1-^{13}C glucose /2- ^{13}C glucose labeled mixed sample.

NaCl, 1 mM DTT pH 8 buffer (GE Healthcare). Monomeric form of MAVS CARD was made by changing buffer pH to 3 to disrupt its self-assembly. For solution state NMR measurements, samples containing 300 µM protein were kept in 50 mM NaH_2PO_4, 1 mM DTT, pH 3 buffer with 0.01% NaN_3 to inhibit bacterial growth. For solid state NMR, different approaches for labeling were used, such as ^{15}N labeling, ^{13}C labeling, ^{15}N-^{13}C uniform labeling, sparse labeling by 1- or 2- ^{13}C glucose. These different labeling strategies were similar in sample preparation by using different combination

of the following chemicals: uniformly ^{13}C labeled glucose, 1- or 2- ^{13}C glucose, unlabeled glucose, ^{15}N labeled ammonium chloride, unlabeled ammonium chloride as the sole C or N source in the growth medium. Protein expression and purification were done in the same way as mentioned above. To distinguish the intra- and inter- molecular constraints, diluted sample, $^{15}N/^{13}C$ mixed sample and 1-^{13}C glucose /2- ^{13}C glucose mixed sample were also prepared. Diluted sample is the mixture of unlabeled monomeric protein with monomeric ^{15}N-^{13}C uniform labeled protein in molar ratio 6 to 1 at pH 3. To make MAVS-$CARD_{1-100}$ form fibril again, the pH was brought back to 7 by adding appropriate amount of 0.5 M NaOH. $^{15}N/^{13}C$ and 1-^{13}C glucose /2- ^{13}C glucose mixed sample were made in the same way with the mixture of ^{15}N labeled protein with ^{13}C labeled protein at equal molar ratio and the mixture of 1-^{13}C glucose labeled protein with 2- ^{13}C glucose labeled protein at equal molar ratio respectively (**Figure 12**).

2.2.3 Protein analyse and characterization

2.2.3.1 SDS polyacrylamide gel electrophoresis

SDS PAGE was used to analyze protein production, purity and molecular weight (Laemmli, 1970). The samples mixed with SDS loading buffer were heated for 5min at 95 °C. 10 ul were loaded per lane on the commercial NuPAGE Novex 4-12% Bis-TRIS Midi Gels (Invitrogen, 1.0 mm x 26 well) and then run as described in the manufacturers' manual. Afterward, the gel was washed with water for few minutes to get rid of SDS and then stained with Instant Blue coomassie dye for 20 min. Distaining of the gel was carried out with water for 5 min.

2.2.3.2 Mass spectrometry and circular dichroism

Mass spectrometry analysis was performed to confirm the size of the produced proteins as well as the correct labeling of the proteins. MALDI-TOF (Matrix Assisted Laser Desorption Ionisation – Time Of Flight) was used for these analyses. The samples for MS analysis were prepared in water. All experiments were carried out by Dr. Manfred Nimtz, Anja Meier and Undine Felgenträger (HZI, Braunschweig)

Circular dichroism (CD) is used to determine the secondary structure contents of a protein. The light wave length for CD is usually within UV range and differential absorption of left and right circularly polarized light by the protein is recorded. Disdistinct absorption characterizes the secondary structure contents. For measurements, the protein samples with 1 mM peptide bonds concentration were applied to 1 mm path length cuvettes. Spectra were recorded using spectropolarimeter JASCO J-810 from far UV range 190 nm to 250 nm at 298K.

2.2.3.3 Pull down assay

Pull down assay are used to study interaction between RIG-I_{1-172} with wt $MAVS_{1-100}$ and different mutants. RIG-I_{1-172} is expressed as fused bait protein with His tagged SUMO* at its N terminal. To carry out the interaction assay, His-SUMO*-RIG-I_{1-172} was mixed with Ni-NTA which bound with His_6 tag. wt $MAVS_{1-100}$, $MAVS_{1-100}$ D23S E26S, $MAVS_{1-100}$ R23S E26S and $MAVS_{1-100}$ R64S R65S mutants were prepared as prey proteins to allow to interact with the bound bait fusion protein in 20mM TRIS, 50mM NaCl, 1mM DTT pH 8 buffer for 1h. Then the non-specifically bound protein were washed by wash buffer for 20CV. Afterward, the complexes were eluted by elution buffer containing 500 mM imidazole and analyzed by SDS-PAGE.

2.2.4 Solution state NMR

2.2.4.1 Data acquisition and processing

All the spectra were collected at 298 K by Brucker Avance 600, 700 NMR spectrometers both equipped with cryo-probes. For backbone and side chain assignment, a series of 2D experiments ^{15}N-HSQC, ^{13}C-HSQC and 3D experiments HNCO, HN(CA)CO, HNCACB, CBCA(CO)NH, HHN-TOCSY (60 ms) and HHC-TOCSY (60 ms) were acquired. HH^{13}C-NOESY-HSQC (80 ms) spectrum was acquired for aromatic sidechain assignment. For distance constraints aliphatic HH^{13}C-NOESY (80 ms), aromatic HH^{13}C-NOESY (80 ms) and HH^{15}N NOESY (80 ms) spectra were recorded. Data were processed with Prosa (Güntert, P et al. 1992) and analyzed with CARA (Keller, 2004). For measurements of relaxation parameters of ^{15}N, couples of spectra were acquired using standard approaches (Boyd et al., 1990; Kay et al., 1989). The spin-spin relaxation constant R2

measurements were done with delays of 17 ms, 34 ms, 51 ms, 85 ms, 119 ms, 153 ms, 187 ms, 237 ms, 288 ms and 356 ms. The R1 measurements were done with delays of 10ms, 30ms, 50ms, 100ms, 200ms, 300ms, 400ms, 800ms, 1000ms, and 1200ms. The detailed acquisition parameters for solution NMR were list in **Table 9**. The relaxation rates and the experimental errors were calculated by following the decay of the height of each H-^{15}N peak from a series of spectra using a build-in option in CcpNmr (Moresco et al., 2011).

Table 9. Acquisition parameters for solution state NMR spectra

Experiment	Acqusition parameters				
	Acqusition time			Temperature	Mixing time
	t1/ms	t2/ms	t3/ms	K	ms
^{15}N-HSQC	122.4	45.1		298	-
^{13}C-HSQC (ali.)	244.7	9.1		298	-
^{13}C-HSQC (aro.)	122.4	42.4		298	-
HNCO	121.7	14.8	20.6	298	-
HN(CA)CO	121.7	13.2	23.6	298	-
HNCACB	243.4	12.3	15.3	298	-
CBCA(CO)NH	121.7	14.8	6.4	298	-
HH^{15}N-TOCSY	113.6	26.3	21.3	298	60
HH^{13}C-TOCSY	113.6	10.3	21.3	298	60
HH^{13}C-NOESY (ali.)	85.2	7.1	17.9	298	80
HH^{13}C-NOESY (aro.)	85.2	11.9	17.9	298	80
HH^{15}N-NOESY	85.2	20.5	17.9	298	80
HSQC-T1[a]	104.4	45.1		298	-
HSQC-T2[b]	104.4	45.1		298	-
^{1}H-^{15}N NOE[c]	91.7	13.2		298	8s(saturation time)

a. HSQC with a delay for inversion recovery

b. HSQC with Carr-Purcell-Meiboom-Gill (CPMG) delay

c. HSQC were recorded in an interleaved manner for NOE and NO NOE where NO NOE is a standard HSQC, NOE is preceded by proton saturation transfer step.

2.2.4.2 Backbone and side chain resonance assignment

Sequence specific backbone resonances were assigned following the standard methodologies mentioned elsewhere (Ahmed et al., 2007). In summary, a suite of in-house written LUA scripts (www.lua.org), were used under the environment of CARA (Keller, 2004) in order to pick peaks and assign spins based on the chemical shift values of resonances in the several spectra. First, HNCO peaks were used to create spin systems based on the unique resonances of each peak due to amide proton and amide nitrogen of the current residue H[i] N[i] and the carbonyl carbon of the previous residue ^{13}C'[i-1]. Following this step, the peaks in ^{1}H-^{15}N HSQC were used to pick carbonyl carbon resonance of the current residue C'[i] along the third dimension of HN(CA)CO. Then picking peaks in the spectrum of CBCA(CO)NH allowed the assignments of both alpha and beta carbons of the previous residue C_α[i-1] and C_β[i-1], and finally HNCACB allowed the identification of the alpha and beta carbons of the current residue C_α[i] and C_β[i] and confirmed the assignment of the previous residue C_α[i-1] and C_β[i-1]. By the end of these steps, all spin systems contained information to identify their amino acid type, as well as their previous amino acid type.

All of these aforementioned peaks' picking and assignments were done as previously mentioned using in-house written LUA scripts, and an extra manual checking of the correctness of the automatic peak-picking and assignments process were essential at the end in order to confirm these assignments. The chemical shift values of each spin system were used to link them together in order to achieve the sequence specific assignments of backbone resonances. Most of the sequence specific connectivity was established using AUTOLINK (Masse and Keller, 2005). The remaining assignments were obtained manually using iterative trials.

Next, we have used the backbone assignments in order to achieve side chain carbon and proton chemical shift assignments. This process started by analyzing a 3D HH^{15}N-TOCSY in order to pick both aliphatic and most of the aromatic proton resonances for each residue. The rest of the aromatics were picked and assigned from HH^{13}C-NOESY in the aromatic region. Once the protons are assigned, we used HH^{13}C-TOCSY in order to assign aliphatic carbon chemical shifts for each residue. The aromatic carbon chemical shifts were assigned by using HH^{13}C(aromatic)-NOESY and H-^{13}C(aromatic)-HSQC spectra. The assignments of both protons and carbons could only be available after the amino acid type assignment which was achieved in the previous step, otherwise complete assignments of these resonances would be much more tedious and more ambiguous.

2.2.4.3 Chemical shift index for secondary structure

Assessment of secondary structure elements in MAVS-CARD was done based on the deviations of the chemical shifts from random-coil values. The chemical shift values for random coil (Wang and Jardetzky, 2002b) were corrected for any sequence dependence through the use of sequence-specific correction factors obtained for a set of Ac-GGXGG-NH_2 peptides in 8 M urea at pH 2.3.(Schwarzinger et al., 2001) We only used the deviations of both C_α and C_β ($\Delta\delta C_\alpha$ and $\Delta\delta C_{\beta)}$, as they showed the strongest dependence on the secondary structure. The deviations of C_α are always in the opposite sense to those of C_β. Here we used chemical shift index that relies on a sum of both effects in order to define in more accurate way any little secondary structure propensity for a given amino acid, namely,

$$CSI = \Delta\delta C^{\alpha} - \Delta\delta C^{\beta} \qquad (1.11)$$

2.2.4.4 Structural calculation

The ^{1}H, ^{13}C and ^{15}N chemical shift assignments and NOESY spectra were used as initial experimental inputs for iterative structure calculation with program UNIO 10 Version 2.0.2 (Guerry and Herrmann, 2012). The NOESY peaks were automatically picked via ATNOS and subsequently assigned by CANDID (Herrmann et al., 2002). The structure calculation was performed using CYANA2.1 (Guntert et al., 1997) with standard simulated annealing schedule. 50 lowest energy conformers consistent with experimental constraints were selected to be refined in explicit water by CNS version 1.2 (Sun et al., 2006). The 20 lowest energy structures in the refinement were selected to represent the final ensemble of structures for MAVS CARD. In the end, the structures were validated on the ICING server (Doreleijers et al., 2012).

2.2.5 Solid state NMR

2.2.5.1 Data acquisition and processing

All NMR measurements were performed on a Bruker Avance III 600 MHz spectrometer operating at a static field of 14.1 T. equipped with a standard 3.2 mm Bruker triple-resonance MAS probe. All spectra were recorded at a sample temperature of 5 ± 1 °C. The pulse sequences were implemented

as recently reported (Shi and Ladizhansky, 2012). Sequential backbone ^{13}C and ^{15}N resonance assignments were achieved using a suite of 3D experiments, namely NCACX, NCOCX, and CONCA experiments. In addition, side chain assignments were completed using a series of ^{13}C-^{13}C and ^{13}C-^{15}N correlation spectra, namely a series of DARR (Takegoshi et al., 2001) experiments with different mixing times and 3D N(CA)CX. Details of experimental setup are given in **Table 10**. All the spectra were processed using Topspin. The backbone dihedral angles were obtained from the C_α and C_β resonances using TALOS+ (Shen et al., 2009). 86 were good predictions and 16 predictions in warning category were not taken since they were located in N,C terminus or loop regions.

Table 10. Acquisition parameters for solid state NMR spectra

Experiment	Acqusition parameters				
	Acqusition time			MAS	Mixing time
	t1/ms	t2/ms	t3/ms	rate/Hz	ms
Uni-CARD DARR	24.5	15.5	-	13.5K	15/50
Uni-CARD DARR	24.5	12.1	-	13.5K	50/100/150/ 250/400
Gado-*Uni-CARD DARR	24.6	15.6	-	13.5K	15
Uni-CARD PDSD	24.6	15.6	-	13.5K	1000
Dilu-CARD DARR	24.5	12.1	-	13.5K	400
Uni-PAR	24.6	15.6	-	13.5K	7
Uni-2D-NCA	18.4	16.0	-	13.5K	-
Uni-2D-NCO	15.8	16.0	-	13.5K	-
2-^{13}C-Glu-NCA	21.5	16.0	-	13.5K	-
Gado*-2-^{13}C-Glu-NCA	21.5	16.0	-	13.5K	
Uni-3D-NCACX	15.3	5	12	13.5K	20/300
Uni-3D-NCOCX	15.3	6.7	12	13.5K	35/250
Uni-3D-CONCA	15.3	6.7	12	13.5K	-
1-^{13}C-Glu CARD PDSD	24.6	15.6	-	13.5K	250/1000
2-^{13}C-Glu CARD PDSD	24.6	15.6	-	13.5K	250/1000
1/2-^{13}C-Glu CARD PDSD	24.6	15.6	-	13.5K	1000
2-^{13}C-Glu CARD -NCA	21.5	16.0	-	13.5K	-
1-^{13}C-Glu CARD PAIN	22.5	16.0	-	13.5K	-
2-^{13}C-Glu CARD PAIN	22.5	15.0	-	13.5K	-

2.2.5.2 Resonance assignment

The sequential walk was achieved by connecting resonances from NCACX, CANCO, NCOCX and 25 ms DARR spectra. Firstly, 2D carbon carbon correlation spectrum was analyzed to find out 'fingerprint' of various types of amino acids such as Thr, Ser, Ile, Ala, Gly, Glu, Gln, Asp, Asn, all of which had easily identified specific patterns. In the sequence of MAVS CARD, both isoluecine and glycine were appeared three times and they were quite unique and well resolved. We took them as the starting point to identify the neighboring residues. Pair-wise residues such as G50 S49, I73 A74, P66 G67, I27 L28 that formed unique doublets can be assigned unambiguously. The left G78, I12 were then identified. For sequence-specific assignment of the residues, the ^{15}N chemical shift of the backbone was needed. In this sense, band-selective heteronuclear NCA and NCO correlation spectra were recorded. The NCA spectrum showed intra-residual information which mean it connected $^{15}N[i]$and $13C_{\alpha}[i]$ from the same residue. The NCO spectrum showed information between $^{15}N[i]$ with previous carbonyl carbon 13C' [i-1]. For sequential assignment walk, NCA and NCO connection were linked to side-chain spins by extending into a 3D experiments. They were achieved by adding one more non-specific 13C homonuclear transferring step and named as NCACX and NCOCX respectively. The resonance information from NCACX and NCOCX were connected by CONCA spectra, which contains the information of $^{13}C'$ [i-1], $^{15}N[i]$and $^{13}C_{\alpha}[i]$. The assignment walk was done using CcpNmr analysis software as mention elsewhere (Stevens et al., 2011; Vranken et al., 2005). To fully assign almost all the side chain resonance, the longer mixing time NCACX (300ms) and DARR spectra (100ms/250ms/400ms) of uniform labeled sample and PDSD spectra (250ms, 1000ms) of sparsely labeled sample spectra were also recorded and used. The 300ms NCACX spectrum was also usded to confirm the assignment because the $^{13}C_{\alpha}[i]$ also exchanged magnetization with its previous residue.

2.2.5.3 Structure determination by solid state NMR

The structure was calculated with ARIA 2.3 (Bardiaux et al., 2008; Rieping et al., 2007) and CNS software (Brunger, 2007). Backbone dihedral angles were predicted with TALOS+ (Shen et al., 2009), and predictions classified as “good” were converted into ϕ and ψ dihedral angle constraints. 86 hydrogen bonds, derived from TALOS secondary structure predictions, were introduced as distance constraints in the structure calculation. The upper bound for distance constraints from

ssNMR cross-peaks was set to 8 Å. Prior to the ARIA calculation, an ensemble of conformer was calculated on the basis of dihedral angle and hydrogen bond constraints only. From this ensemble, the maximum observed distance (*Dm*) of the non-long-range unambiguous ssNMR constraints was checked, and if *Dm* was lower than 8 Å, the upper bound of the distance was set to *Dm*. For the final ARIA iteration, 100 conformers were calculated and the 15 lowest energy structures were refined in a shell of water molecules (Linge et al., 2003).

2.2.6 Negative staining electron microscopy

2.2.6.1 Sample preparation

Negative staining sample preparation, according to the Valentine et al. (Valentine et. al.), has been applied to visualize purified MAVS by mean dose energy-filtered transmission electron microscopy (EFTEM). Formation of protein fibers and sample homogeneity were thus checked. Molecules from working solution of finally 100 µg/mL protein concentration were adsorbed to a 40 nm ultra-thin carbon foil by floating partially from a 1 x 1 mm^2 carbon-coated piece of mica for 30 sec. at ambient temperature. After blotting with filter paper (Schleicher-Schuell, Dassel, Germany) the carbon foil was transferred shortly to saturated uranyl acetate solute pH 4.0, was picked up with a reticulum foil coated 300 mesh copper grid (Lünsdorf et. al. 1986), the stain bulk was blotted and finally air-dried.

2.2.6.2 Energy-filtered transmission electron microscopy

Ultra-structural analysis of negatively stained molecules was performed with an EF-TEM (Libra 120 plus, Zeiss, Oberkochen, Germany), equipped with a bottom-mount cooled 2048x2048 CCD-camera (SharpEye, Tröndle, Moorenweis, Germany). Elastic bright-field images, free from objective astigmatism, (energy-slit width: 10 eV; objective aperture: 60 µm; beam current: 2 µA; acceleration voltage: 120 kV) were recorded at x40000 and/or x50000 magnification under close to low-dose conditions within the range of Gaussian and Scherzer focus (150 nm to 250 nm under focus).

2.2.7 Electron microscopy images processing

The micrographs we got were 16 bit images in TIFF format. The first step was to box the straight linear helical particles using 'helixboxer' command from EMAN2 package. Power spectra were generated using the 'helix_automation' script from IHRSR++. The power spectra can be used either to extract symmetry information or to sort the boxed images. Because, in many cases, the samples may contain fibers failed to diffract. The power spectrum provides a convenient way to discard these fibers. The average power spectra could be generated by using command 'avg_pw2' to average all non-repeated power spectra.

The vertical line passes through the middle of the pattern is called meridian. The horizontal line in the center of the pattern is called equator. The other layer lines encoding the symmetry information of the helices are distributed on both sides of the meridian with mirror symmetry. The distance from the layer line to the meridian is the radius in Fourier space (R), which is used to calculate the Bessel order n according to the following equation

$$2\pi Rr = 1.1\,|n| + 0.9 \qquad (1.12)$$

r is the radius of the filament which can be estimated from EM images. The distance from the layer line to the equator is related to layer line number *l*, which is used for the helical selection rule (1.10). With Bessel order *n* and layer line number *l*, rise per subunit Δz and rotation per subunit $\Delta\Phi$ could be calculated with the equation from M.F.Moody's book.

$$\Delta z = |\,(N\, l_1^{-1} l_2^{-1}) / (n_1\, l_1^{-1} - n_2\, l_2^{-1})| \qquad (1.13)$$

$$P = |\,(N\, l_1^{-1} l_2^{-1}) / (k_1\, l_1^{-1} - k_2\, l_2^{-1})| \qquad (1.14)$$

$$k_1 n_2 - k_2 n_1 = \mathrm{N} \qquad (1.15)$$

$$\Delta\Phi = 360^{o}\,(\Delta z / P) \qquad (1.16)$$

N is the order of the rotation axis, which equals to the highest common factor of n_1 and n_2. *P* is the pitch of the helix which means the width of one complete helix turn along the axis.

2.2.8 In vivo cell experiment

2.2.8.1 Electroporation

NIH3T3 IFN-b-tGFP cells were cultured as monolayer in DMEM medium supplemented with 10% fetal bovine serum and G418. Cells are cultured at 37 °C and an atmosphere of 5% CO_2 in air to reach 80% confluence. Treat growing cells with trypsin/EDTA till the cells were suspended and removed the trypsoin/EDTA by centrifugation at 300g. The cell pellet were washed with PBS at ambient temperature. Re-centrifuged the cells and resuspended them in PBS to 5×10^6 cells/ml. Then took 800 μl of the cell suspension into a 0.4 cm electroporation cuvet, adding MAVS fiber and control protein with final concentration 100 μg/ml. Electroporated the cells with 0.3 kv and 125 uF capacitance. The time constant should be between 12 to 22 ms under this condition. The cell suspension were then removed to a flask containing pre-warmed medium. The fluorescence of tGFP was checked by fluorescence microscopy and flow cytometer 16 hour after electroporation.

3 Results

3.1 Monomeric wt MAVS-CARD structure and dynamics by solution NMR

3.1.1 Expression and purification of MAVS and RIG-I

The plasmid pETRO2.11 containing N-terminal 6×His tagged SUMO* fused before the synthesized gene of MAVS $CARD_{1-100}$, $MAVS_{1-175}$, $MAVS_{1-500}$ and $RIG\text{-}I_{1-172}$ were transformed into T7 cells and the protein was expressed in CN040 medium. MAVS-$CARD_{1-100}$ forms filaments when over-expressed in *E.coli* (Hou et al., 2011). Likewise, the fusion protein 6×His-SUMO*-MAVS $CARD_{1-100}$ was found as high-molecular weight structures in the supernatant after cell lysis. Attempts to purify these filaments by immobilized metal ion affinity chromatography failed, because it bound only weakly to Ni-NTA in 20 mM TRIS, 50 mM NaCl, 2 mM DTT, pH 8.0 buffer. However, the filaments of MAVS-$CARD_{1-100}$ could be separated from most other soluble components in the supernatant by ultracentrifugation at 100,000 g for 2 h. In the pellet the purity of MAVS-$CARD_{1-100}$ was more than 70%, estimated from SDS-PAGE gel. This pellet was dissolved again in 20mM TRIS,

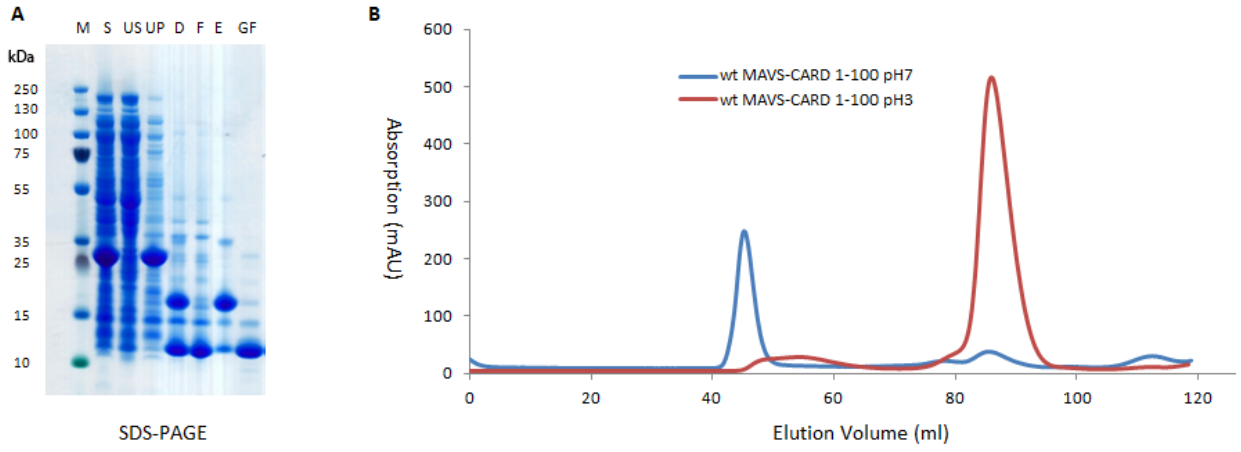

Figure 13. SDS-PAGE of the purification of MAVS-$CARD_{1-100}$: A. M: marker; S: supernatant; US: supernatant after ultra-centrifugation; UP: Pellet after ultra-centrifugation; F: Flow through fraction of Ni-NTA column; E: Elution fraction; GF: fraction from gel filtration column by superdex 200. **B.** Gel filtration chromatography of MAVS-$CARD_{1-100}$ at pH 3 and 7 in 50mM sodium phosphate buffer.

50mM NaCl, 2mM DTT pH 8 buffer and cleaved with SUMO* protease at 4 °C overnight. The sample was then applied to Ni-NTA column to remove the His-tagged SUMO*. Since small pieces of cell membranes and nucleic acids could also be precipitated together with MAVS-CARD$_{1-100}$ in the ultracentrifugation step, another gel filtration step was applied to get pure MAVS-CARD$_{1-100}$. Since MAVS-CARD$_{1-100}$ formed filaments, it was eluted in the void volume (**Figure 13**). The protein expression level and fractions in each purification step were analyzed by SDS-PAGE gel (**Figure 13**). In contrast to MAVS-CARD$_{1-100}$, the other three proteins MAVS$_{1-175}$, MAVS$_{1-500}$ and RIG-I$_{1-172}$ could bind on Ni-NTA column, and they were purified by immobilized metal ion affinity chromatography through a standard protocol. The molecular weight of these three proteins were checked by gel filtration using superdex 200 column. The result revealed MAVS$_{1-175}$, MAVS$_{1-500}$ formed smaller aggregation comparing to MAVS-CARD$_{1-100}$ and RIG-I$_{1-172}$ contained both high molecular weight and low molecular weight components (**Figure 14).**

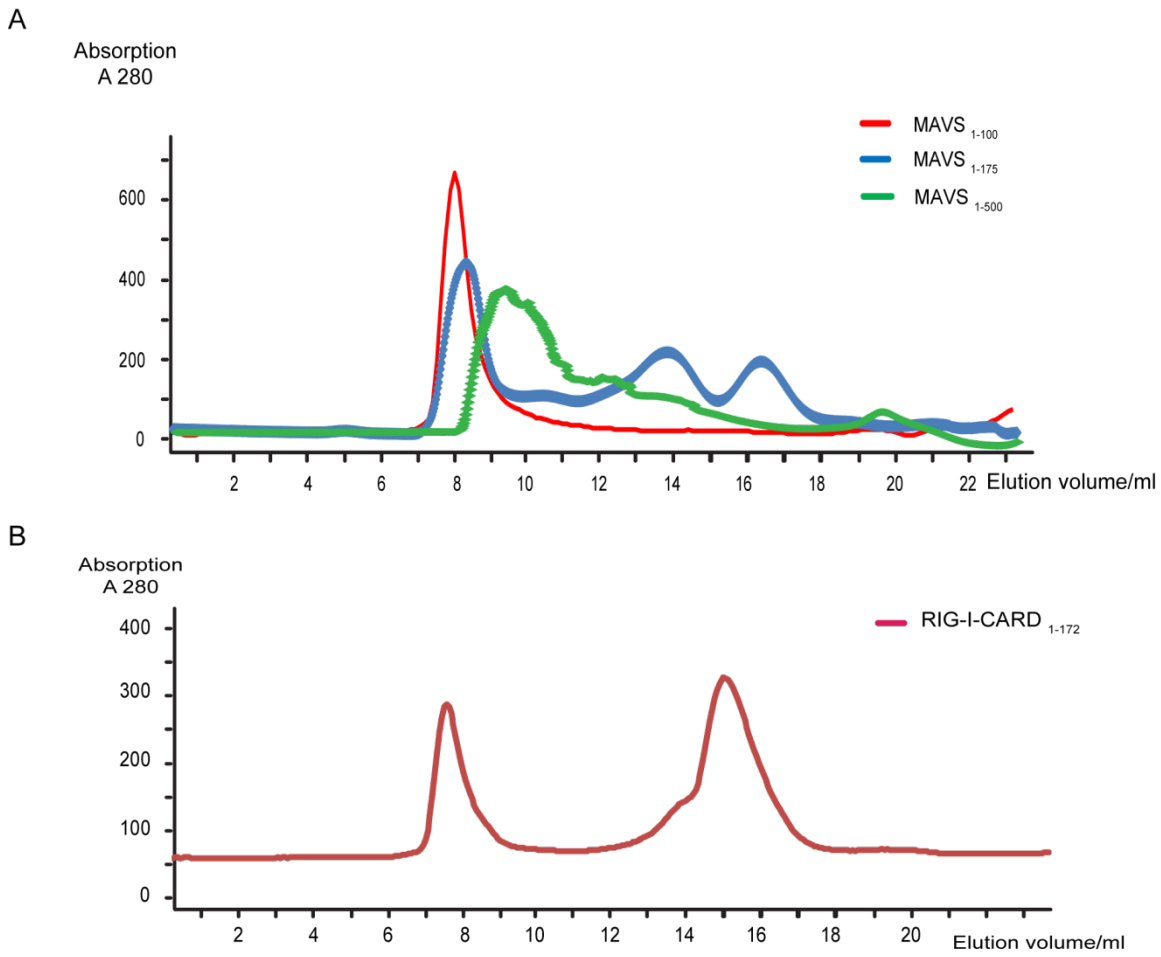

Figure 14: Gel filtration chromatography of MAVS$_{1-175}$, MAVS$_{1-500}$ and RIG-I$_{1-172}$ by Superdex 200. A. MAVS $_{1-100}$ (CARD alone) contained only high molecular weight components. MAVS $_{1-175}$ contained both high and lower molecular

weight components. MAVS $_{1\text{-}500}$ had lower molecular weight components compared to CARD alone. **B.** Gel filtration of RIG-I$_{1\text{-}172}$ by superdex 200.

3.1.2 Morphology of MAVS-CARD$_{1\text{-}100}$ in different buffers

Freshly purified MAVS-CARD$_{1\text{-}100}$ was analysed by transmission electron microscopy (TEM) for investigation of its ultrastructure. TEM analysis revealed that MAVS-CARD$_{1\text{-}100}$ formed regularly ordered long fibers with a diameter of approximately 8.2 ± 0.2 nm (**Figure 15**).

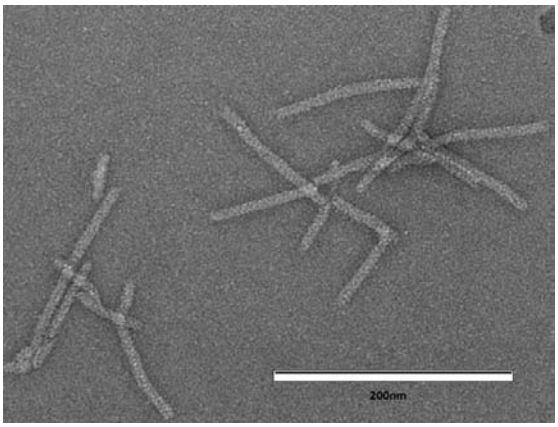

Figure 15. The TEM image of MAVS-CARD$_{1\text{-}100}$: MAVS-CARD$_{1\text{-}100}$ assembles into filaments with diameter around 8.2 ±0.2 nm.

Previous studies on CARDs (Park et al., 2007a) showed that the interactions between CARDs are mainly mediated by electrostatic interaction. So, buffer screening of MAVS-CARD$_{1\text{-}100}$ was done to find out if there is a condition to convert it from a filamentous state to a monomeric state. ^{15}N labeled MAVS-CARD$_{1\text{-}100}$ in different buffer conditions was monitored by NMR (**Figure 16**). A series of 2D-NMR spectra revealed that MAVS filaments could be disrupted by low pH buffer. In 50mM pH 7 phosphate buffer, MAVS-CARD$_{1\text{-}100}$ was in a fibrillar state. Due to the slow tumbling of the protein filaments, there was no signal in the HSQC spectrum (**Figure 16 A**). The theoretical isoelectric point of MAVS-CARD$_{1\text{-}100}$ is 5.72, and indeed it was found that MAVS-CARD$_{1\text{-}100}$ precipitated in pH 5 and 6 buffer. In 50mM sodium acetate pH 4 buffer, MAVS-CARD$_{1\text{-}100}$ filaments were disrupted, resulting in lower molecular weight proteins (**Figure 16 B**). However, the HSQC spectrum showed inhomogeneous, low intensity peaks, indicating that MAVS-CARD$_{1\text{-}100}$ is in

multiple, oligomeric states in this buffer (**Figure 16 B**). The HSQC spectrum of MAVS-CARD$_{1-100}$ in 50mM phosphate buffer pH=3 showed high intensity and homogeneous peaks, indicating that the protein was monomeric (**Figure 16 C**), which was confirmed by gel filtration of MAVS-CARD$_{1-100}$ in the same buffer condition (**Figure 13 B**).

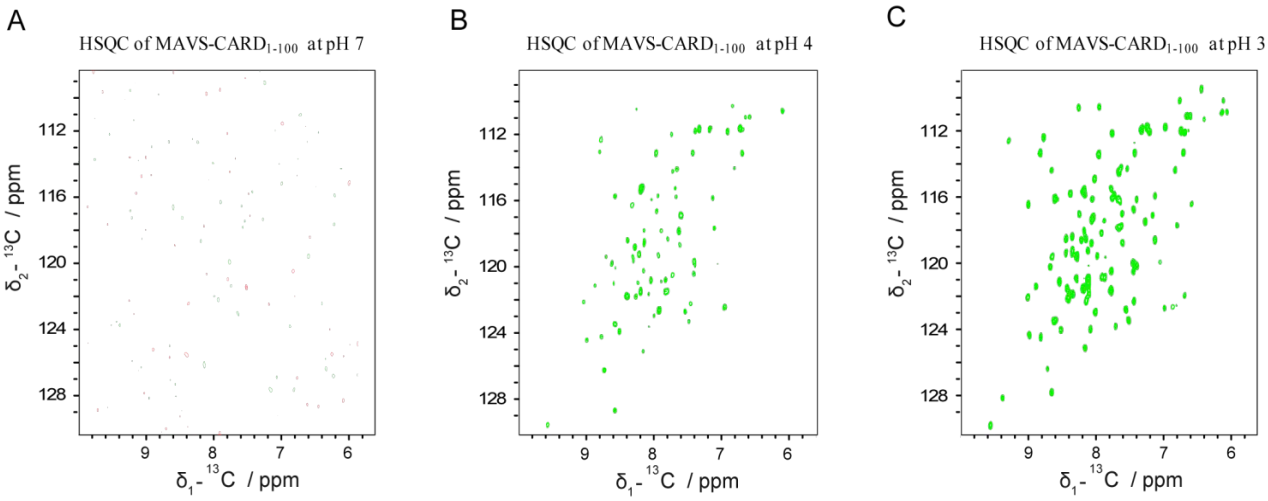

Figure 16. Buffer screening of MAVS-CARD$_{1-100}$ monitored by NMR: A. In pH 7 buffer, MAVS-CARD$_{1-100}$ formed high molecular weight filaments. There is only noise on the HSQC spectrum due to its short T2 relaxation time. **B.** In pH 4 buffer, MAVS-CARD$_{1-100}$ started to convert to low molecular weight complexes. In pH 3 buffer, MAVS-CARD$_{1-100}$ showed good dispersion of $^{1}H^{15}N$ resonances in both dimensions of the HSQC spectrum.

3.1.3 Secondary structure and intrinsic fluorescence of MAVS-CARD$_{1-100}$

The secondary structure of MAVS-CARD$_{1-100}$ were determined using circular dichroism (CD) spectroscopy in the far-UV spectral region (190-250 nm). The minima at 222 nm and 208 nm and a maximum at 190 nm are indicative of a mostly alpha helical protein (**Figure 17 A**). The folding status of MAVS-CARD$_{1-100}$ at pH 3.0 and pH 7.0 were studied by measuring intrinsic fluorescence. The intrinsic fluorescence intensity as well as the emission maxima of the aromatic residues tyrosine and tryptophan (λmax) are strongly dependent on the local molecular environment. Consequently intrinsic protein fluorescence could be used to probe the folding status of the protein's tertiary structure. Typically, the quantum yields of aromatic amino acids decrease in well-folded proteins, and emission maxima below 340 nm are indicative of aromatic residues buried in the hydrophobic

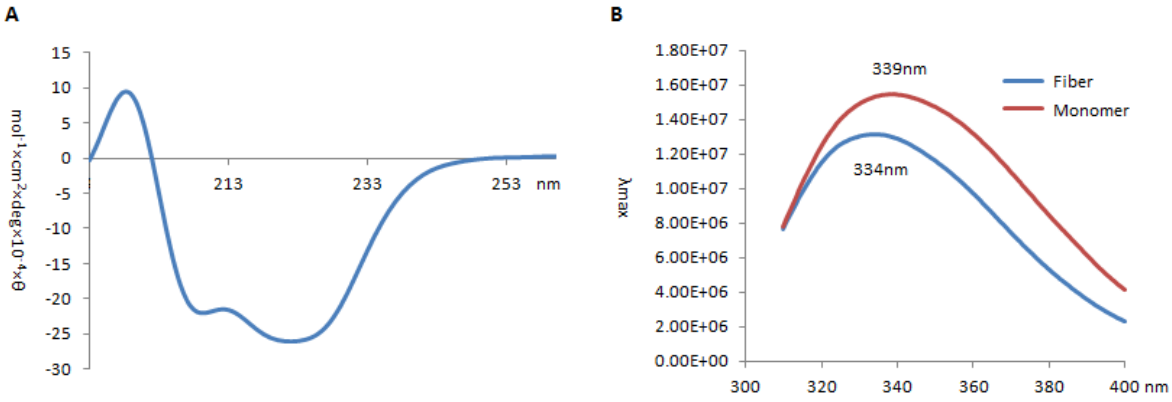

Figure 17. Biophysical characterization of MAVS-CARD$_{1-100}$: **A.** CD spectroscopy of MAVS-CARD$_{1-100}$ reveals typical alpha helix secondary structure. **B.** Intrinsic fluorescence of MAVS-CARD$_{1-100}$ indicates that the filamentous form is less accessible and possibly more stable than the monomer.

core of the protein. MAVS-CARD$_{1-100}$ appears to be well folded both in the monomeric and in the fibrillar state. The slightly lower fluorescence intensity, and the lower emission maximum indicate that the aromatic side chains become more buried upon fibril formation, and might be involved in a more stable structure (**Figure 17 B**).

3.1.4 Chemical shift assignment of MAVS-CARD$_{1-100}$ at pH 3.0

Since MAVS-CARD$_{1-100}$ was monomeric at pH 3.0, uniformly ^{13}C, ^{15}N-labled MAVS-CARD$_{1-100}$ in 50 mM sodium phosphate buffer with 7% D_2O and 1 mM DTT at pH 3.0 was used for measuring multidimensional NMR spectra for structure determination and analysis of the dynamic properties. To achieve backbone and side chain resonance assignment, 2D $^{1}H^{15}N$ HSQC, $^{1}H^{13}C$ HSQC and 3D HNCO, HN(CA)CO, CBCA(CO)NH, HNCACB, HH^{15}N-TOCSY, HH^{13}C-TOCSY and HH^{15}N-NOESY were recorded and analyzed with the software CARA (Keller, 2004). $^{1}H^{15}N$ signals of 96 residues were observed in 2D ^{1}H-^{15}N HSQC spectra, while the 4 proline residues did not show a $^{1}H^{15}N$ signal (**Figure 19**). The good dispersion of the $^{1}H^{15}N$ resonances in both dimensions indicated a well-folded protein and facilitated the assignments of the spectra. Regarding the sequential backbone assignment, the sequential connectivity information C_α[i], C_β[i], C'[i] C_α[i-1], C_β[i-1], C'[i-1] could be obtained from HNCACB and HN(CA)CO spectra and the C_α[i-1], C_β[i-1], C'[i-1] information were confirmed by CBCA(CO)NH and HNCO spectra. To obtain distance

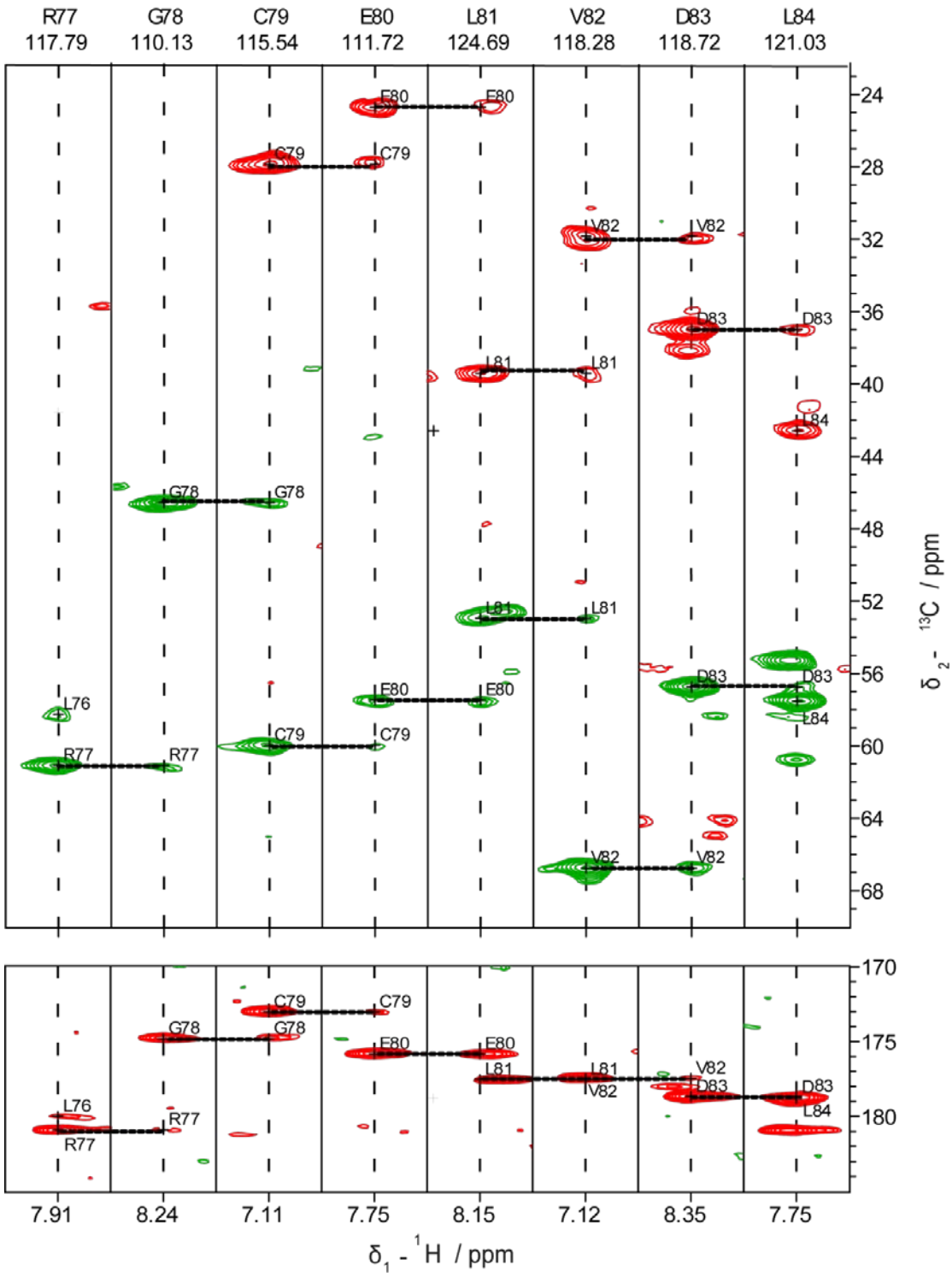

Figure 18. strip plots extracted from HNCACB and HN(CA)CO for backbone assignment of MAVS-CARD$_{1\text{-}100}$. The sequential connections were obtained from HNCACB (top) and HN(CA)CO (bottom) through correlating of C_α[i], C_β[i], C'[i] C_α[i-1], C_β[i-1], C'[i-1] information. CA, CO resonances are shown in red, and CB resonances are shown in green.

constraints for structure calculation, HH^{15}N-NOESY and HH^{13}C-NOESY were employed. An Autolink LUA script was used to build the connection for most of the residues (Masse and Keller, 2005). The side chain assignment was completed from HH^{15}N-TOCSY, HH^{13}C-TOCSY spectra manually. As an example, parts of the HNCACB and HN(CA)CO spectra are shown in **Figure 18** to exemplify the assignments and connectivity. All of the 102 residues including the N-terminal GS residues that remained after protease cleavage were assigned unambiguously. The assignment is shown in **Figure 19**. Resonances of backbone amide groups were labeled in black. Side chains of Arg, Asn, Trp, Gln were labeled in blue. Arg side chains are folded into the spectrum and the peaks were indicated in green.

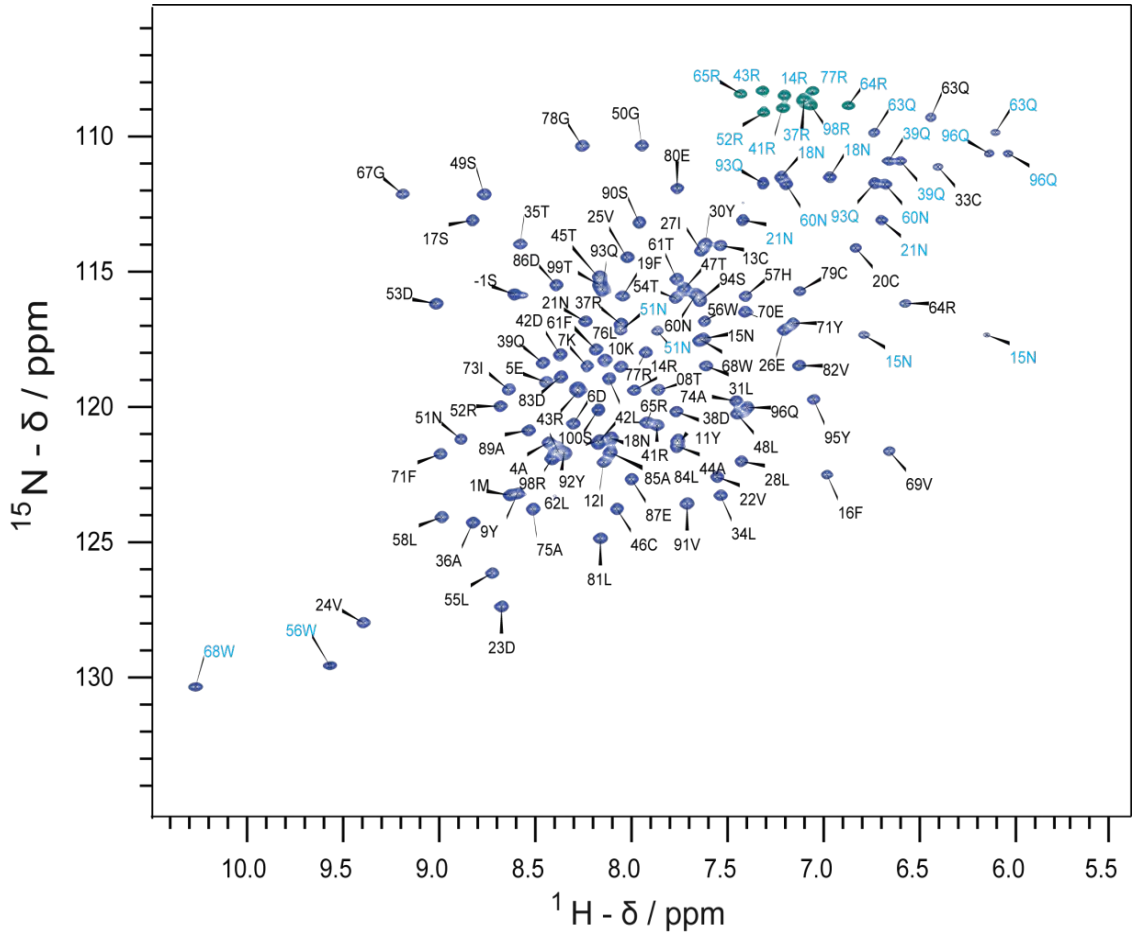

Figure 19. Assigned 2D ^{1}H-^{15}N HSQC spectrum of ^{13}C, ^{15}N labeled MAVS-CARD$_{1\text{-}100}$: 300µm MAVS-CARD$_{1\text{-}100}$ in 50 mM NaH_2PO_4, 1 mM DTT, pH 3.0 buffer (H_2O/D_2O 15:1 v/v) acquired at a proton resonance frequency of 600.28 MHz. Resonances of backbone amide groups are labeled in black. Side chains of Arg, Asn, Trp, Gln are labeled in blue. Arg side chains are folded into the spectrum and the peaks are indicated in green.

3.1.5 Structure calculation of monomeric MAVS-CARD$_{1-100}$ by solution NMR

The structure of monomeric MAVS-CARD$_{1-100}$ was calculated under help from Dr. Johannes Spehr with the software UNIO (Guerry and Herrmann, 2012). 1064 NOE distance constraints, and 505 dihedral angle restraints were used. The 20 lowest energy structures (**Figure 20 B**) had a root mean square deviation (RMSD) of 1.52 ± 0.17 Å for backbone heavy atoms of secondary structure elements (**Table 11**). The solution NMR structure of MAVS-CARD$_{1-100}$ showed the typical topology of the death domain (DD) superfamily, comprising six α-helices arranged in a Greek-Key motif (**Figure 20 A**). The ensemble of the best 20 structure were shown in. Helix1 extended from residue 2 to 18. Helix2 (residue 24 to 27) was connected to Helix 1 and Helix 3 by long loops. The other four helices were Helix3 (residue 36 to 49), Helix4 (residue 51 to 64), Helix 5 (residue 68 to 79) and Helix 6 (82 to 95). The overall structure was similar to the crystal structure of MAVS CARD fused to Maltose Binding Protein (Potter et al., 2008), but the loops on both sides of helix 2 were longer, suggesting high flexibility of helix 2 in solution (**Figure 20 C**).

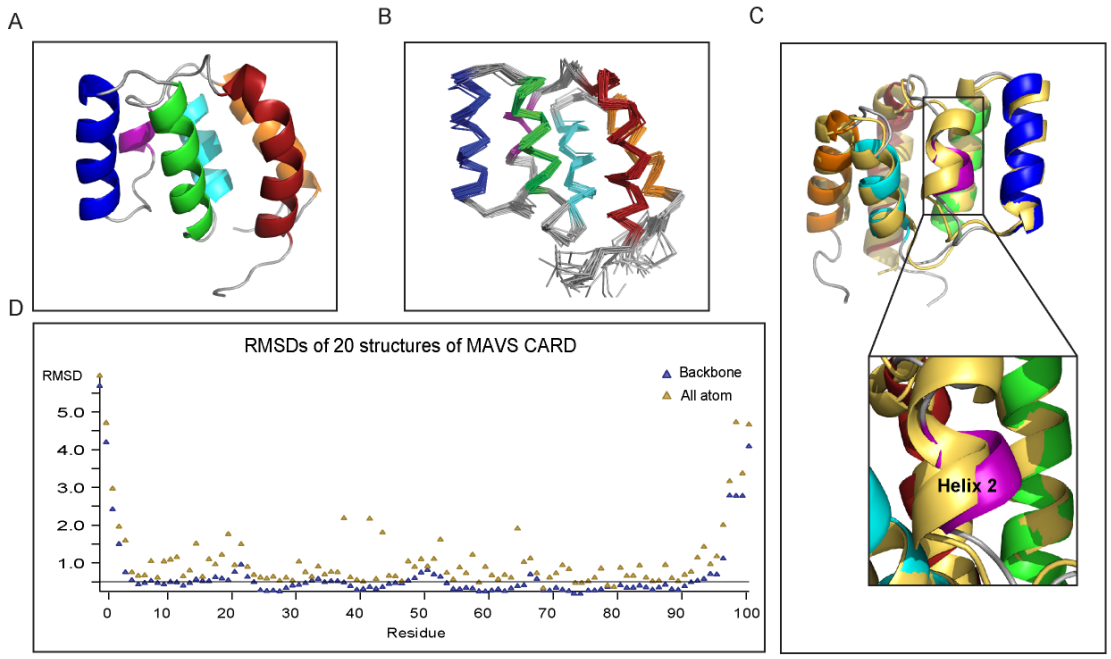

Figure 20. Structure calculation of MAVS-CARD$_{1-100}$ by solution NMR: A. cartoon diagram of the structure; **B.** Backbone superimposition of the 20 lowest-energy structures of monomeric MAVS-CARD$_{1-100}$.; **C.** Overlay of solution NMR structure with the previously published MBP-fused MAVS-CARD$_{1-93}$ crystal structure (Potter et al., 2008). The crystal structure is in yellow. **D.** R.M.S.D values per residue for the ensemble of the 20 best structures, given for the backbone (blue) and all atoms (green), respectively.

Table 11. Summary of structural constraints and structure statistics

Parameter All structures[a]	result	
	calculation	refinement
NOE distances total	1064	
intraresiduel (| i-j | = 0)	454 (42.7%)	
sequential (| i-j | = 1)	302 (28.4%)	
medium range (| i-j | < 4)	165 (15.5%)	
long range (| i-j | > 5)	143 (13.4%)	
dihedral angles	505	
residual target function (Å²)	9.8 ± 0.4	
residual violations		
NOE >0.1 Å		5 ± 2
dihedral angles > 2.0°		19 ± 3
CNS-energy (PARALLHDG-force field) (kcal/mol)		
total		-3745.2 ± 54.8
van der Waals		-461.6 ± 15.3
electrostatic		-3968.1 ± 58.4
RMSD (Å) [b]		
backbone (4-96) Å		0.98 ± 0.12
heavy atom (4-96) Å		1.52 ± 0.17
Ramachandran statistic [c]		
favored %		83.4
allowed %		14.7
generous allowed %		1.1
disallowed %		0.8

[a] Average values and standard deviations over the 20 conformers
[b] Average root mean square deviation (RMSD) over the 20 conformers' atomic coordinates with respect to the average structure.
[c] Percentage of residues in the Ramachandran plot regions determined by CING (Doreleijers et al., 2012)

3.1.6 Surface characteristics

CARDs are a subclass of the death domain superfamily. Structural homology search by the DALI server indicates that the structure of MAVS-CARD$_{1-100}$ is closest to RIG-I CARD, Apaf-1 and

prodomain of procaspapse-9 with a Z score lager than 11 (Holm and Rosenstrom, 2010). The structure similarity to RIG-I CARD domain will be studied and discussed in **section 4.2**. The other two proteins, Apaf-1 and prodomain of procaspapse-9, form a complex mainly via electrostatic interaction, where the negatively charged patch Asp 27, Glu40 and Glu42 of Apaf-1 interacts with the positively charged patch Arg10, 11, 13, 52, and 56 of procaspase-9 prodomain. This interaction surface is also reinforced by H-bond interaction and van der Waals interaction (Vaughn et al., 1999). The surface electrostatic potential plot of MAVS-CARD$_{1\text{-}100}$ also shows two charged patches. Lys7, Lys10, Arg14, Arg37, Arg41, His57, Arg64 and Arg65 comprise the positively charged surface and residues Asp23, Glu26, Glu80, Asp83 and Asp86 comprise the negatively charged surface (**Figure 21**). The observation that MAVS-CARD$_{1\text{-}100}$ polymer could be disrupted to form monomers at pH 3.0 also indicates the electrostatic interactions might be a driving factor for the MAVS-CARD$_{1\text{-}100}$ self-assembly into filaments. Other oligomeric CARD domain complexes revealed there are three types of interaction involved in CARD-CARD homotypic association. Type I interaction is from the loop between helix 2 and 3 with helix 1 and 4. Type II interaction is from the loop between helix 4

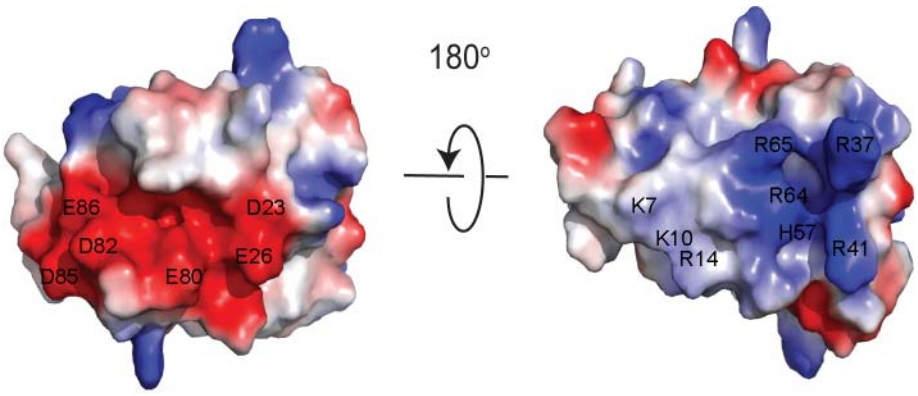

Figure 21. Plots of surface electrostatic potential. Positive and negative charges were colored blue and red respectively. The neutral residues were colored white. The surface plots were generated by PyMOL (The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC).

and 5 with the loop between helix 5 and 6. Type III interaction is from helix 3 with the loops between helices 1 and 2 between helices 3 and 4 (Ferrao and Wu, 2012; Lin et al., 2010; Park et al., 2007b). As mentioned in **section 3.1.5,** the major difference of the NMR structure to the previously determined crystal structure is the loops on both sides of helix 2, which are longer in the NMR

structure. This might indicate a larger flexibility of these two loops, facilitating MAVS-CARD$_{1-100}$ complex formation through the type I and III interactions (**Figure 45**).

3.2 MAVS-CARD$_{1-100}$ mutants and their interactions with RIG-I CARD

3.2.1 Mutants of MAVS-CARD$_{1-100}$

The electrostatic surface plot of the MAVS-CARD$_{1-100}$ structure (**Figure 21**) and the phenomenon that the fibrils of MAVS-CARD$_{1-100}$ can be disrupted in lower pH buffer strongly indicated the charged amino acids are involved in the assembly of the MAVS-CARD$_{1-100}$ filament. Therefore, several of these surface charged amino acids were mutated to see if they could disrupt the formation

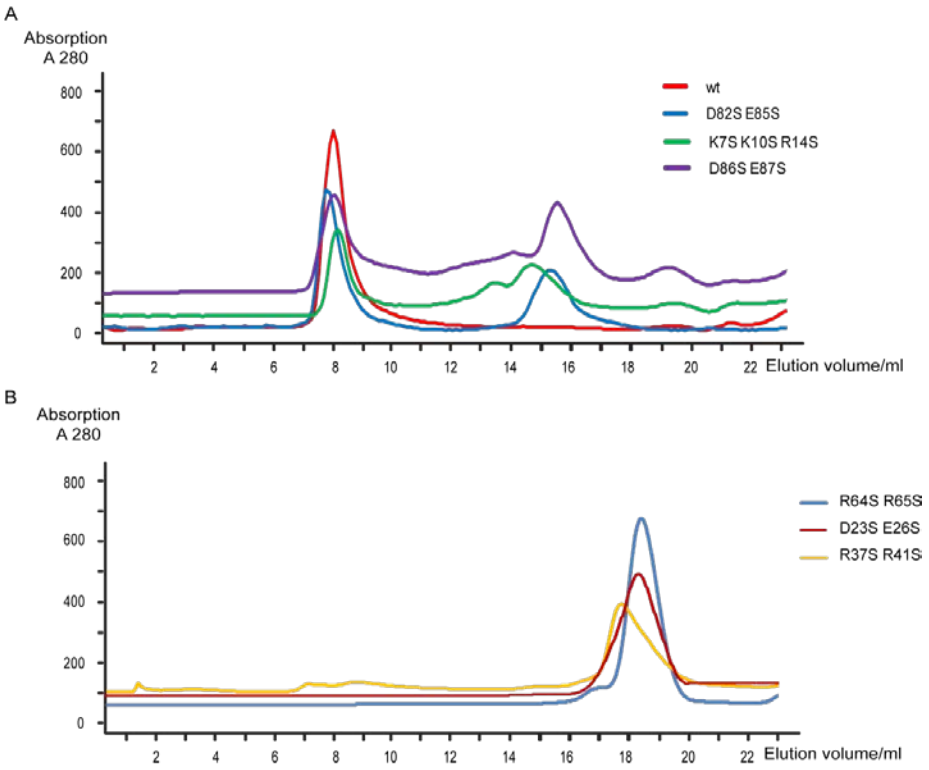

Figure 22. Size exclusion chromatography of MAVS-CARD$_{1-100}$ mutants: Chromatograms of MAVS-CARD$_{1-100}$ mutants by Superdex 200. Absorptions at 280 nm were plotted against the elution volumes.

of the filament. Interestingly, three double mutants were found to completely abolish the ability to self-assemble at neutral pH. The size and assembly status of all these mutants at pH 7 in phosphate

buffer were checked by gel filtration using a Superdex 200 column (**Figure 22**). The double mutants D23S E26S, R64S R65S and R37S R41S were eluted after 18 ml, corresponding to a molecular weight of 13 kDa, implying that these mutants were in a monomeric state. In contrast, the wt MAVS-$CARD_{1\text{-}100}$ was eluted in the void volume, meaning it was present in oligomeric states with molecular weights larger than 440 kDa. Compared to the wt MAVS-$CARD_{1\text{-}100,}$ the other mutants K7S K10S R14S, E80S D83S and D86S E87S had a second elution peak at a volume of 15 ml, which might be the position of a dimer, indicating that these mutants partially reduced the ability to self-assemble.

3.2.2 Backbone resonance assignment of the monomeric mutant D23S E26S at pH 7.0

To further study the dynamic properties of monomeric MAVS-$CARD_{1\text{-}100}$, and to gain site-specific information on its interaction with the tandem CARDs of RIG-I by NMR, HSQC spectra of the three monomeric mutants D23S E25S, R64S R65S and R37S R41S were recorded. The wide dispersion of $^{1}H^{15}N$ resonances from 6.0 to 10.3 ppm and narrow peak widths in both dimensions indicated all proteins were very well folded (**Figure 23**). This suggests that the mutagenesis that caused MAVS-

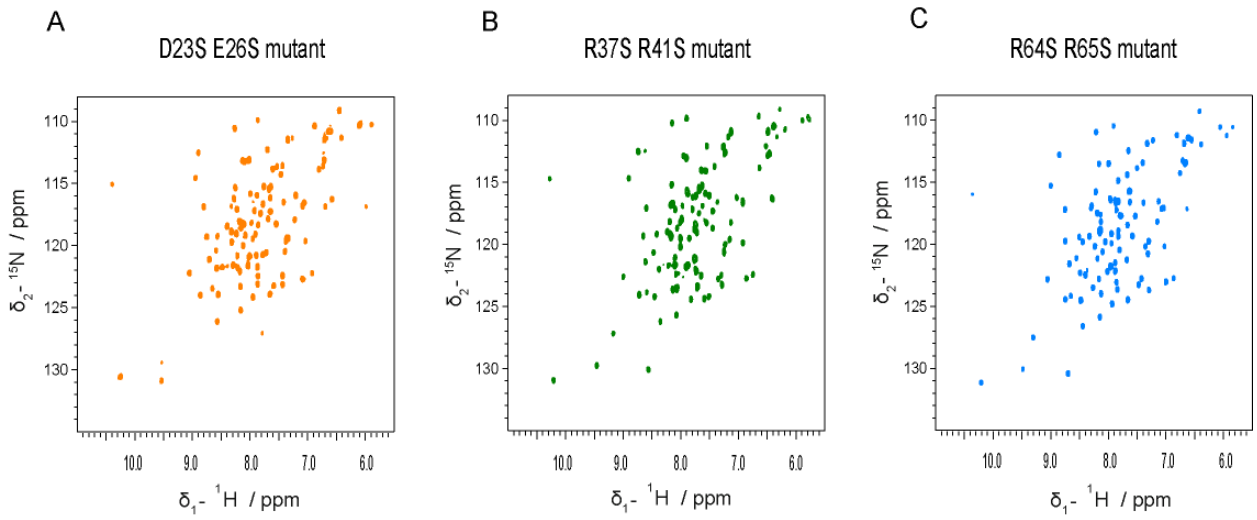

Figure 23. HSQC spectra of the monomeric mutants in 50mM phosphate buffer at pH 7.0: A. HSQC spectrum of D23S E26S double mutant, **B.** HSQC spectrum of R37S R41S double mutant, **C.** HSQC spectrum of R64S R65S double mutant**.**

$CARD_{1\text{-}100}$ to remain monomeric did not affect the folding of the protein. It is thus unlikely that loss of function of monomeric mutants might be caused by a misfolding of the protein.

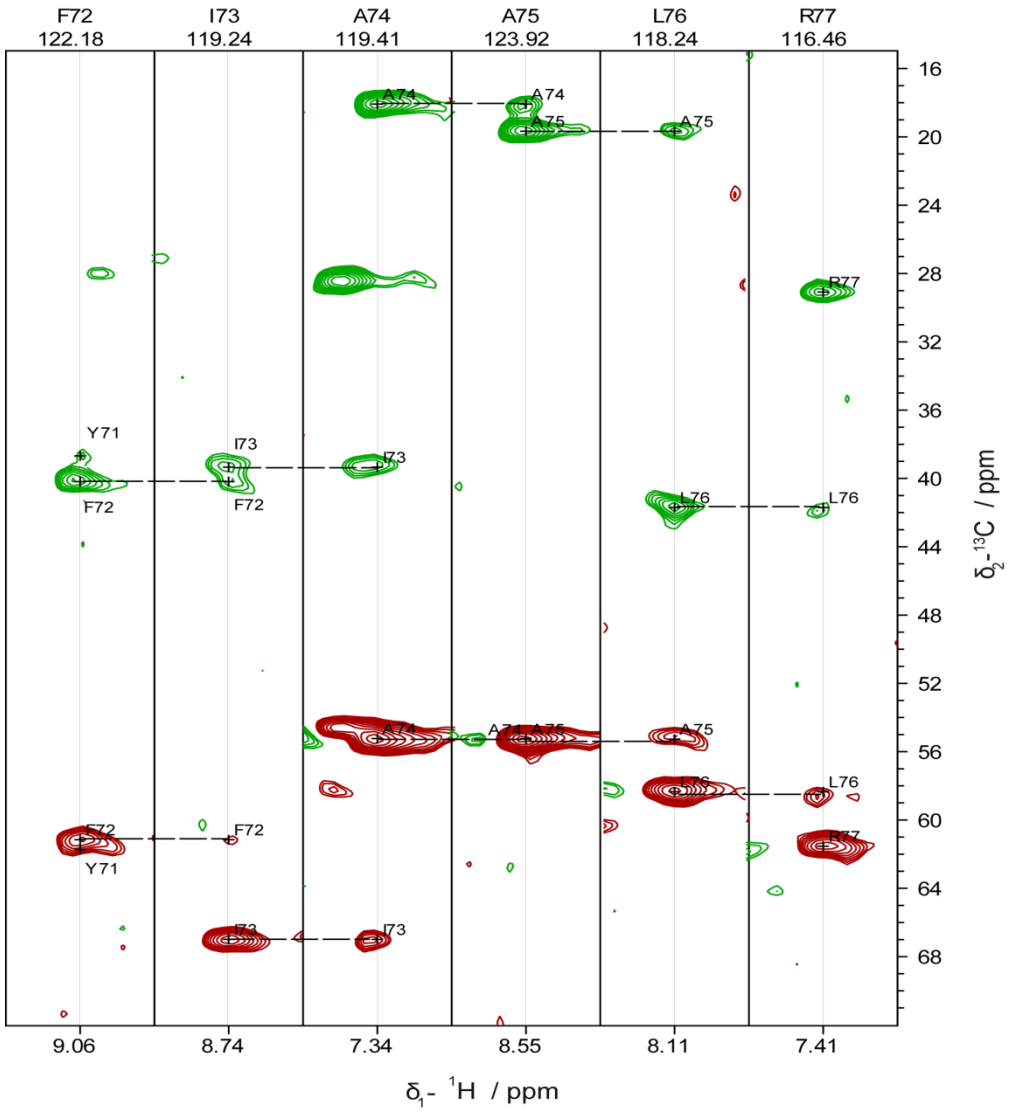

Figure 24. Strip plots extract from HNCACB and HN(CA)CO for backbone assignment of the D23S E26S mutant: The sequential connections were obtained from HNCACB through correlating of $C_\alpha[i]$, $C_\beta[i]$, $C_\alpha[i\text{-}1]$, $C_\beta[i\text{-}1]$, information. CA resonances are colored in red and CB resonances in green.

Because several resonances shifted their position due to the different pH, the backbone resonance assignment for the D23S E26S double mutant in 50mM phosphate buffer pH 7.0 was done as described in **section 2.2.4.2**. Strip plots from the HNCACB spectrum are shown in **Figure 24**.

3.2.3 Dynamics of MAVS-CARD$_{1-100}$

To fully understand the function of MAVS-CARD$_{1-100}$ at physiological pH, a set of ^{15}N relaxation experiments were performed on the mutant D23S E26S to quantify the backbone dynamics on the

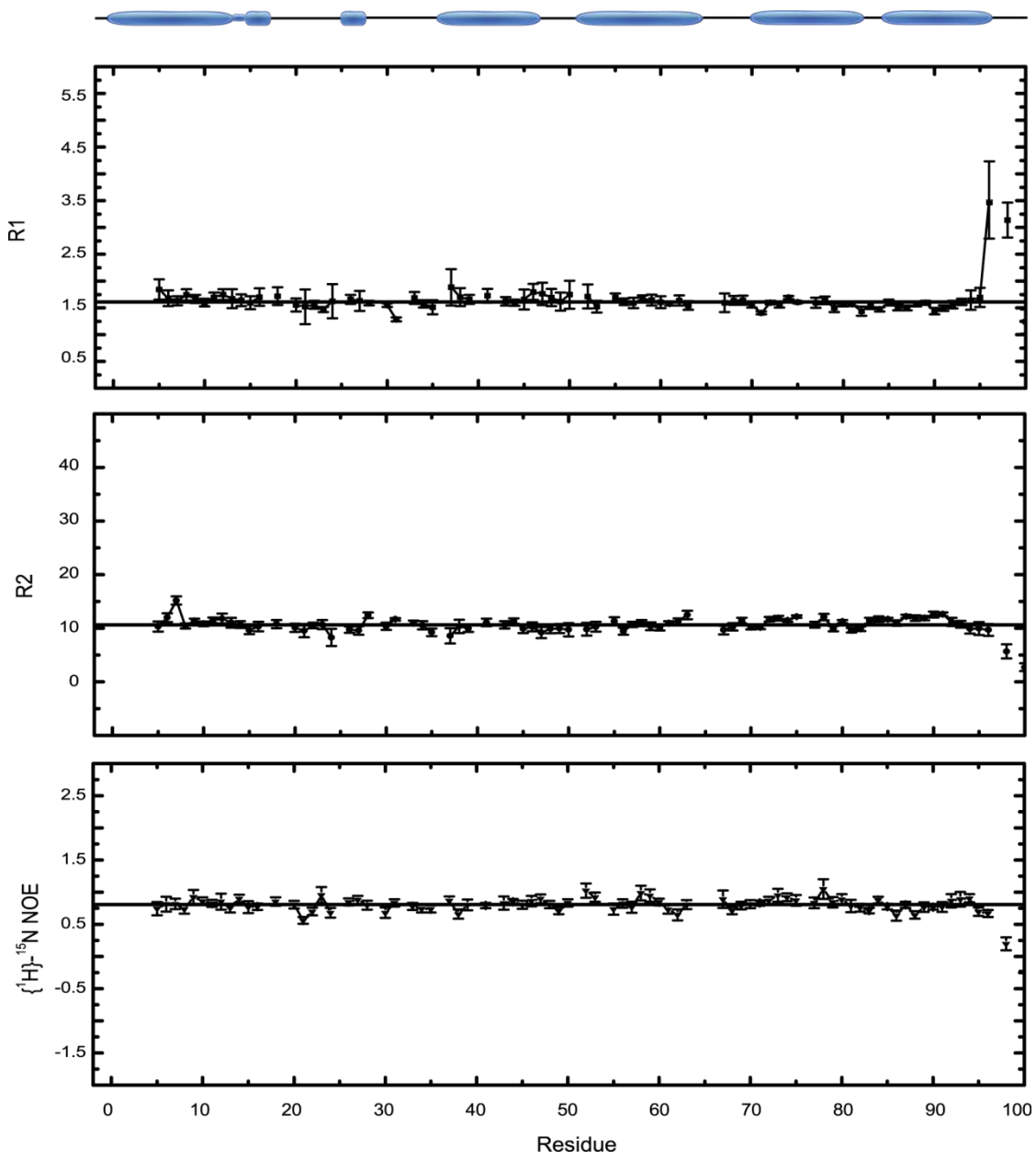

Figure 25. Backbone relaxation parameters R_1, R_2 of MAVS-CARD$_{1-100}$: The secondary structures are indicated on the top of the chart.

picoseconds-to-nanoseconds time scale as well as on the microsecond-to-millisecond time scale. Longitudinal R_1, transverse R_2 relaxation and heteronuclear ^{1}H-^{15}N NOEs were measured to determine motions within a molecule on the atomic level. 86 assigned residues were analyzed and the peak heights were used for curve fitting by CCPN (Vranken et al., 2005) to avoid effects of partially overlapping resonances, which occurred in particular for the residues K10 with L76 and N21 with T45. The R_1 measurements showed that the relaxation rates ranged from 1.22s^{-1} to 3.36 s^{-1} with an average value of 1.61s^{-1}, which is the typical value for a folded protein of similar size. The R2 values ranged from 2.7 s^{-1} to 15.1 s^{-1} with an average value of 10.64 s^{-1} (**Figure 25**). For the C terminus, the R_2 values were lower, indicating random coil structure at the end of the MAVS-$CARD_{1-100}$. The overall rotational correlation time τ_c was 6.6 ns, estimated from relaxation data of backbone R_2/R_1. This value indicates that the mutant D23S E26S was monomeric at physiological pH. For heteronuclear ^{1}H-^{15}N NOE, most of the residues exhibited positive NOE values, except for few residues at the C terminus. Surprisingly, all loops between the helices displayed NOE values larger than 0.7, similar to those found in the helices, indicating very restricted dynamics for the whole molecule (**Figure 25**).

3.2.4 Reduced spectral density mapping

The relaxation data of MAVS-$CARD_{1-100}$ was interpreted by two ways. The first was the reduced spectral density function approach (Farrow et al., 1995; Peng and Wagner, 1992) and the second was model free analysis (Lipari and Szabo, 1982a, b). The calculated values of the reduced spectral density functions $J_{(\omega H)}$, $J_{(\omega N)}$, and $J_{(\omega 0)}$ are shown in **Figure 26**. An inspection of the data shown in **Figure 26** indicated that, except for the terminal regions, the whole protein was behaving like a rigid body with no differential motions in different secondary structural elements. This is apparent from the almost invariable values of all spectral density functions for all residues except the terminal fragments.

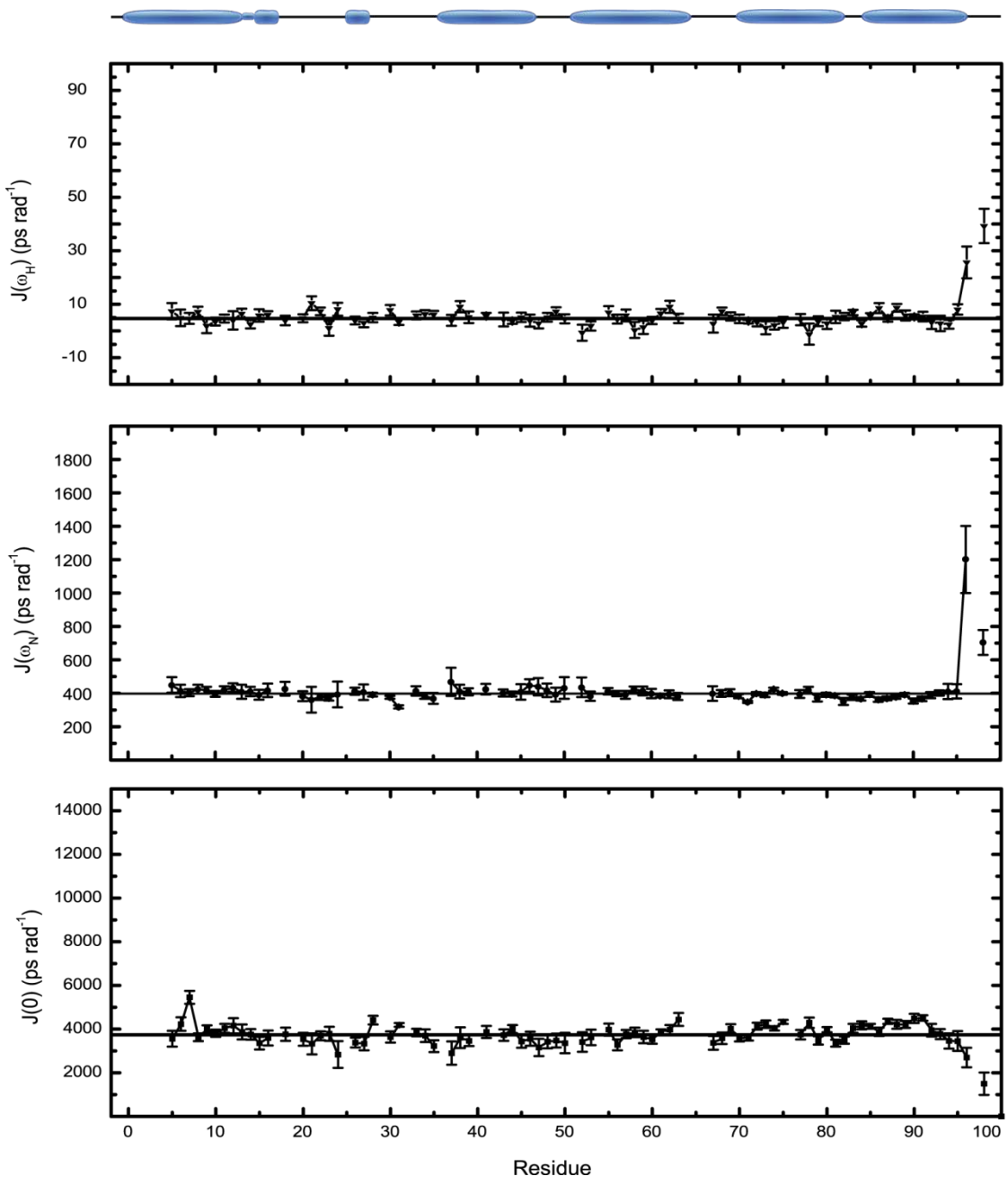

Figure 26. Reduced spectral density function of MAVS-CARD$_{1\text{-}100}$: The secondary structural elements are indicated on the top of the chart.

3.2.5 Model free analysis

A more detailed view of the different local dynamic parameters along the sequence could be obtained by more extensive analysis of the relaxation parameters using model-free analysis. In this

project, I used the program TENSOR2 (Dosset et al., 2000), which used the model-free approach from Lipari and Szabo (Lipari and Szabo, 1982a, b) for determination of the internal mobility parameters. The total correlation time τ_c was calculated by the MODELFREE built-in grid search option. Two rotational diffusion models (isotropic and anisotropic) of the protein were evaluated and the anisotropic diffusion model was found to fit the NMR relaxation data better. A total correlation time of 7.88 ns was used by the algorithm. Once the overall diffusional behavior of the protein had

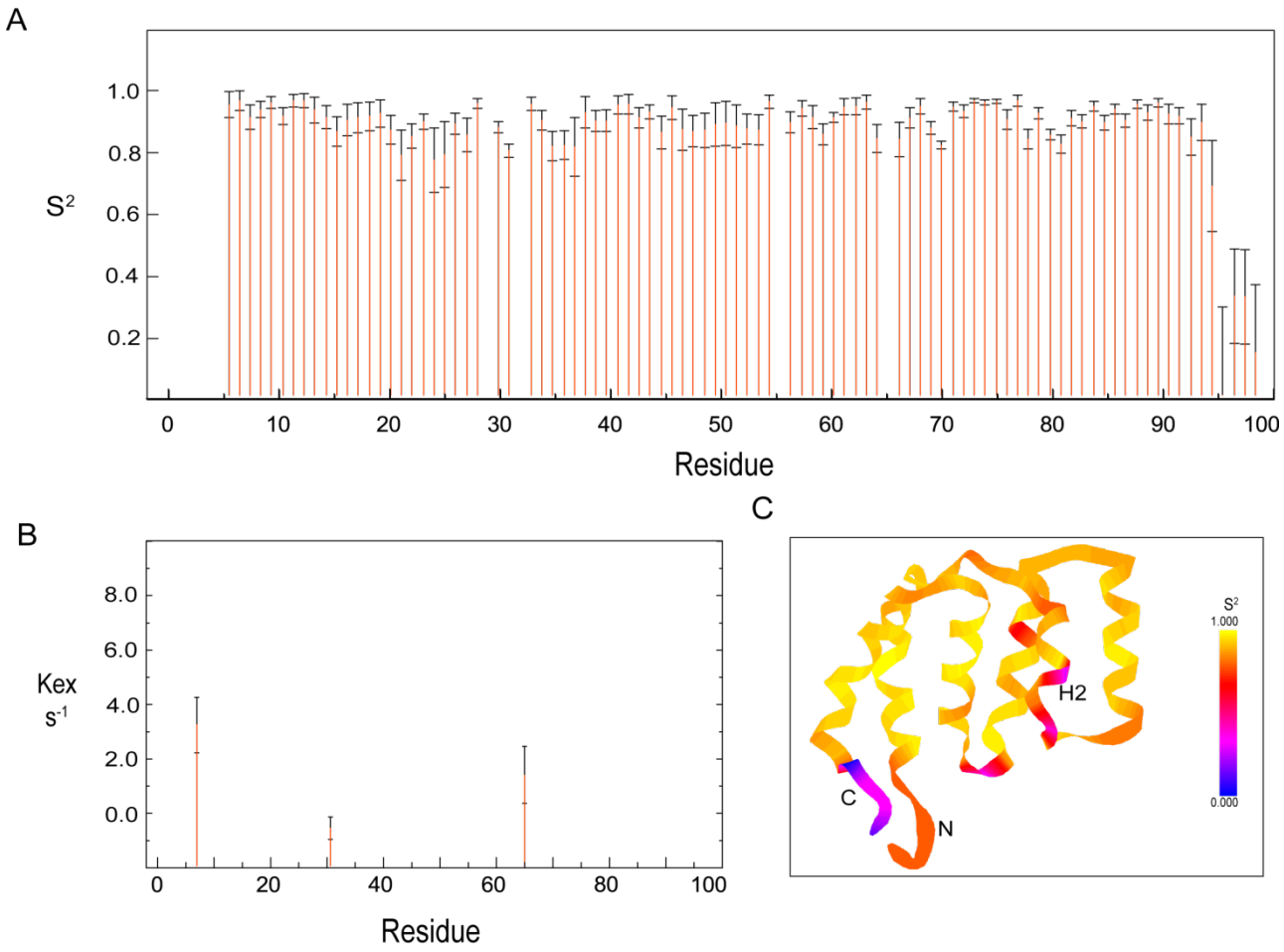

Figure 27. Model free analysis results for D23S E28S mutant: A. S^2 **B.** R_{ex}. For both S2 and $R_{ex,}$ the errors were determined from 1000 Monte Carlo simulations of Brent's implementation of Powel's method for multidimensional minimization. **C.** The parameter S^2 was plotted on the structure of MAVS-$CARD_{1-100}$.

been determined, the program would perform Lipari and Szabo type analysis for the internal mobility. The model selection strategy from Mandel et al. (Mandel et al., 1995) was applied to find a model that can best explain the experimental data. In the end, the internal dynamic properties of the

protein were parameterized as residue-specific measurement of the generalized order parameters (S^2) describing the generalized motion of each H-N vector, and a factor R_{ex} indicating exchange processes on the milli- to microsecond time-scale. The most important parameter to assess fast internal motion was S^2. A value of 1 would indicate total restriction, whereas a value of 0 would indicate no restriction of the internal motion of the N-H bond. **Figure 27 A** shows that the generalized order parameter S^2 was around 0.9 for most of the residues. The dynamic properties of MAVS-$CARD_{1-100}$ revealed no evidence of a contiguous polypeptide chain with increased internal motion, except at the extreme chain termini and the loop between helix 2 and helix 3, which had significantly smaller S^2 values of 0.7 to 0.8. The S^2 values were plotted on the structure of MAVS-$CARD_{1-100}$ with the scale bar from 0 to 1. The pink colored regions were the disordered parts (**Figure 27 C**). While S^2 provided a measurement of rapid internal motions comparing to the overall rotational diffusion, the parameter R_{ex} represents frequency motions for some residues. It identifies residues experiencing conformational or chemical exchange on a microsecond-to-millisecond time-scale, which is the time scale for many biological events such as ligand association and dissociation rates (Mukherjee et al., 2006). The residues K7, R65 were observed to have significant R_{ex} values (>1.5 s^{-1}), indicating there might be conformational or chemical exchange. L31 was also observed to have an R_{ex} value of 1.26. Since only two single residues had significant exchange rate, an unambiguous description of their R_{ex} terms to a particular conformational exchange was not trustable. However, in the following section, these two residues were found to be involved in type I and III interaction surfaces, respectively.

3.2.6 Interaction of MAVS-$CARD_{1-100}$ mutants with RIG I CARD domain

To understand how MAVS-$CARD_{1-100}$ is activated by interaction with the tandem CARD domains of RIG-I and to map the interaction surface, His-SUMO* tagged RIG-I $CARD_{1-172}$ was first immobilized on Ni-NTA beads and then mixed with three different monomeric MAVS-$CARD_{1-100}$ mutants in order to test for co-elution of the two interaction partners. **Figure 28** shows that both positive charged amino acids mutants (R64S R65S and R37S R41S double mutants) lost the ability to interact with RIG-I $CARD_{1-172}$, while the negative charge amino acid mutant D23S E26S and wt MAVS-$CARD_{1-100}$ co-eluted with His-SUMO* tagged RIG-I $CARD_{1-172}$, indicating that they formed a stable complex .

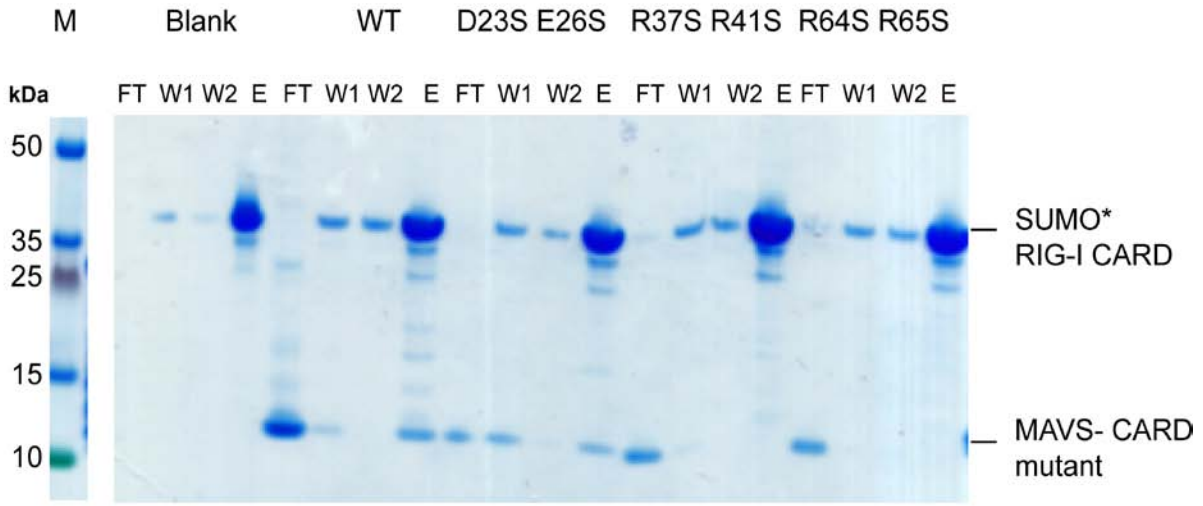

Figure 28. Pull down assay of RIG-I CARD$_{1-172}$ with MAVS-CARD$_{1-100}$ mutants: 6×His-SUMO* tagged RIG-I CARD$_{1-172}$ was immobilized on Ni-NTA beads, and then mixed with wt MAVS-CARD$_{1-100}$ and its mutants. Unspecifically bound protein was washed with 20 column volumes washing buffer containing 30 mM imidazole. In the blank group 20 mM TRIS, 50 mM NaCl, 1 mM DTT pH 8.0 buffer was used instead of MAVS-CARD$_{1-100}$ as negative control.

Since residues R37 R41 R64 and R65 were all in the positively charged patch (**Figure 22**), the mutagenesis on this patch might hit the interaction surface between MAVS-CARD$_{1-100}$ and RIG-I CARD$_{1-172}$. However, detailed site specific information on the interaction surface was still missing. To fill this gap, a ^{15}N labelled D23S E26S MAVS-CARD$_{1-100}$ mutant was titrated with increasing amounts of unlabelled RIG-I CARD$_{1-172}$ without His-SUMO* tag. 2D ^{1}H-^{15}N HSQC spectra of D23S E26S mutant were acquired at 298K with protein/protein molar ratios of 1:0, 1:0.5, 1:1, 1:2. The signal intensities of many resonances decreased with increasing RIG-I CARD concentrations, which indicated that the corresponding amino acids were involved in a chemical exchange of MAVS-CARD$_{1-100}$ mutant with RIG-I CARD (**Figure 29**). The rate of the exchange in NMR experiments is usually classified into three time scales: slow, intermediate and fast. It depends on the value of the chemical shift difference between bound and unbound states relative to the exchange rate constant, K_{ex}. If the chemical shift difference $\Delta\omega = |\upsilon_{bound} - \upsilon_{unbound}|$ between the two states is in the same order as the chemical exchange rate K_{ex}, it is called intermediate rate exchange. In this case, both resonances for bound and unbound state would disappear because of severe line broadening

(Palmer et al., 2001). This is known as chemical exchange line broadening effect (Cavanagh et al., 2007).

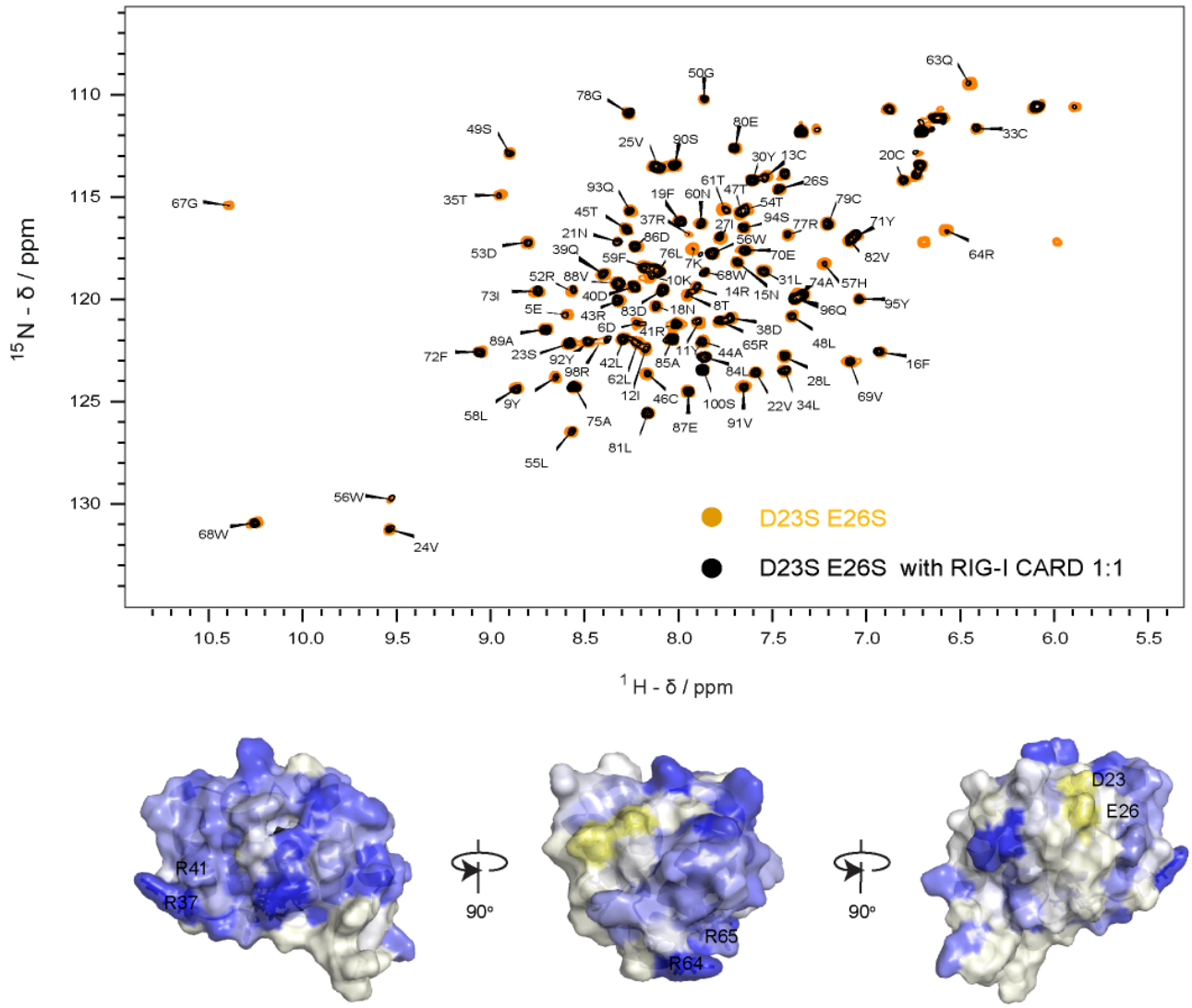

Figure 29. Interaction of D23S E26S mutant with RIG-I CARD$_{1-172}$: A. ^{1}H-^{15}N HSQC spectra were recorded for D23S E26S MAVS-CARD$_{1-100}$ (uniformly ^{15}N labelled, 100 uM) in absence (orange) or presence of unlabeled RIG-I CARD$_{1-172}$ (blank) at molar ratio 1:1. **B.** The residues in D23S E26S that exhibited increasing loss of signal intensity are shown in increasing shades of blue on the surface of MAVS-CARD$_{1-100}$. The residues that were mutated for the pull down assay are marked on the structure.

In addition to the residue specific intermediate exchange upon binding of RIG-I CARD$_{1-172}$ with MAVS-CARD$_{1-100}$ mutant, the effective molecular weight of the complex became larger. Thus the transverse relaxation rates increased, resulting in reduced intensity of all the peaks in the HSQC

experiment. However, The residues with the most significantly decreased intensities were identified and plotted on the surface of MAVS-$CARD_{1-100}$ structure (**Figure 29**). Most signals identified were from the positively charged patch. This result corresponded very well with the pull down assay. Nevertheless, there were other residues from different parts of the molecule also affected by binding of RIG-I $CARD_{1-172}$, which indicated there could be more than one interaction surface. This would be in agreement with the three types interaction surface between CARD domains discussed in **section 1.1.6**.

3.3 Fibrillar MAVS-$CARD_{1-100}$ structure by solid state NMR and EM

To understand the structural basis for MAVS-$CARD_{1-100}$ filament formation and if it had similar interaction surfaces between each subunit as RIG-I $CARD_{1-172}$ binding with it, we applied solid state NMR and EM to study MAVS-$CARD_{1-100}$ filament structure. Both techniques don't rely on fast molecular tumbling and availability of high-quality crystals, hence serving as promising methods for studying the structure of large protein assemblies.

3.3.1 Seeding ability of the *in-vitro* expressed MAVS-$CARD_{1-100}$

To test if the recombinant MAVS-$CARD_{1-100}$ from *E.coli* could still activate the expression of interferon, we used electroporation to introduce MAVS-$CARD_{1-100}$ filament with and without treatment of sonification, as well as the control protein SUMO* into a GFP-fused interferon report cell line (Rand et al., 2012). The SUMO* was expressed together with MAVS-$CARD_{1-100}$ as a fusion protein. After the purification of the fused protein, SUMO* was cleaved by SUMO* protease and used as the negative control since it was in the same condition as MAVS-$CARD_{1-100}$. **Figure 30** showed both MAVS-$CARD_{1-100}$ groups (with and without the treatment of sonification) could activate the immune response and up-regulate the expression of interferon. In contrast, in the negative control group, there was no expression of interferon. Interestingly, the MAVS-$CARD_{1-100}$ treated with sonification before electroporation induced two times more cells expressing interferon than the untreated MAVS-$CARD_{1-100}$, which might be because the sonification broke long MAVS-$CARD_{1-100}$ fibers into smaller fragments, facilitating their transfection into the report cell line. This

result indicated that the *E.coli* expressed MAVS-CARDs filament had bioactivity that could auto-catalytically trigger endogenous MAVS to initiate the immune response.

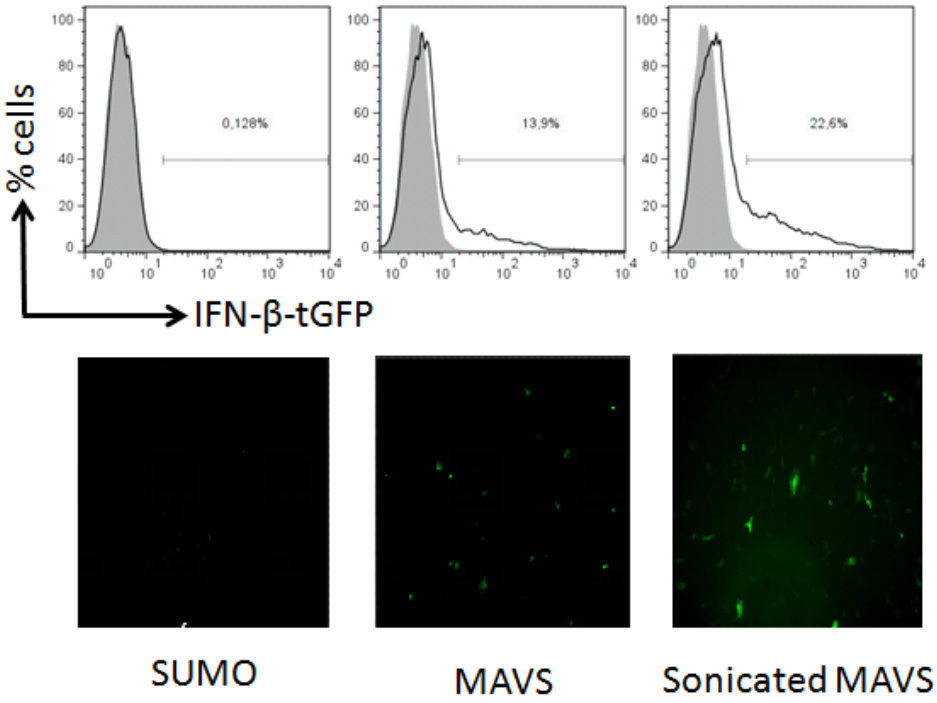

Figure 30. Recombinant MAVS-CARD$_{1-100}$ filaments purified from *E. coli* induced immune response: IFN-β–GFP reporter cells were checked 12h after electroporation of recombinant fibrillar MAVS-CARD$_{1-100}$. The green cells indicated the immune response was triggered and interferon-β was expressed. The level of the expressed interferon was quantified by flow cytometry.

3.3.2 Resonance assignments of the MAVS-CARD$_{1-100}$ filament by solid state NMR

Upon activation, MAVS-CARD mediates the autocatalytic assembly of full length MAVS to aggregate on the mitochondrial surface into a fibrillar state. The filament could then induce inactive MAVS to assemble into a similar structure to rapidly amplify the immune signalling (Hou et al., 2011). To better understand the structural basis for the MAVS activation, an atomic-resolution structure of MAVS-CARD$_{1-100}$ filament would be required. In general, filamentous proteins are unlikely to be crystallized and their high molecular weight renders them incompatible to be studied by solution NMR. However, such regularly ordered filamentous proteins are ideally suited for structural studies by solid state NMR. Since this technique does not rely on fast molecular tumbling,

it has no molecular weight limitation for the studied protein complex. It has been shown that highly ordered assembled protein samples greatly reduce the inhomogeneous broadening of the NMR signals, leading to higher resolution spectra. In this section, the complete assignment of ^{13}C and ^{15}N chemical shifts of fibrillar MAVS-$CARD_{1-100}$ is described, as well as the determination of secondary structural elements of fibrillar MAVS-$CARD_{1-100}$ from the C_α and C_β chemical shifts. In addition, the chemical shifts are also compared with the corresponding chemical shifts of the monomeric state MAVS mutant D25S E28S which were measured in the same buffer condition as fibrillar MAVS-$CARD_{1-100}$ to predict the interaction surface between CARD domains based on the chemical shift perturbation.

3.3.2.1 Spectra quality

Filamentous MAVS-$CARD_{1-100}$ also consists of 102 residues, of which the first two residues are glycine and serine remaining after cleavage of the SUMO* tag. The filament preparation of MAVS-$CARD_{1-100}$ gives highly resolved lines. **Figure 31** shows a 1D ^{13}C spectrum of MAVS-$CARD_{1-100}$ along with its sequence. It was evident that the sample attained a high degree of microscopic order and homogeneity, leading to narrow line widths. This was consistent with the relaxation data (**Figure 23**) that monomeric MAVS-CARD was extremely rigid even for the loops between the helices. This overall rigidity likely contributed to the homogeneity and narrowed the line widths. The highly ordered state and homogeneity of MAVS fibers enabled us to collect high resolution 2D and 3D correlation spectra with excellent signal to noise ratio in reasonable times. To get even higher resolution for the bulky aromatic side chains and get longer distance constraints for structure determination in later stages, the MAVS-$CARD_{1-100}$ filament was also expressed in media with 1-^{13}C glucose and 2-^{13}C glucose as the only carbon source to get sparsely labeled protein. This sparse labeling approach (Castellani et al., 2002; Hong, 1999; Loquet et al., 2011) removed the one bond C-C dipolar and J couplings, resulting in smaller line widths. The 1D spectra of 1- /2- ^{13}C glucose sparsely labeled sample are also shown in **Figure 31**. The line widths of the sparsely labeled samples were about by 1.5~1.8 times smaller compared to the uniformly labeled sample, leading to a resolution comparable to that observed for the HET-s prion domain (Ritter et al., 2005; Wasmer et al., 2008) and nanocrystalline GB1 (Franks et al., 2008), which are considered to be the benchmarks for ssNMR spectral quality.

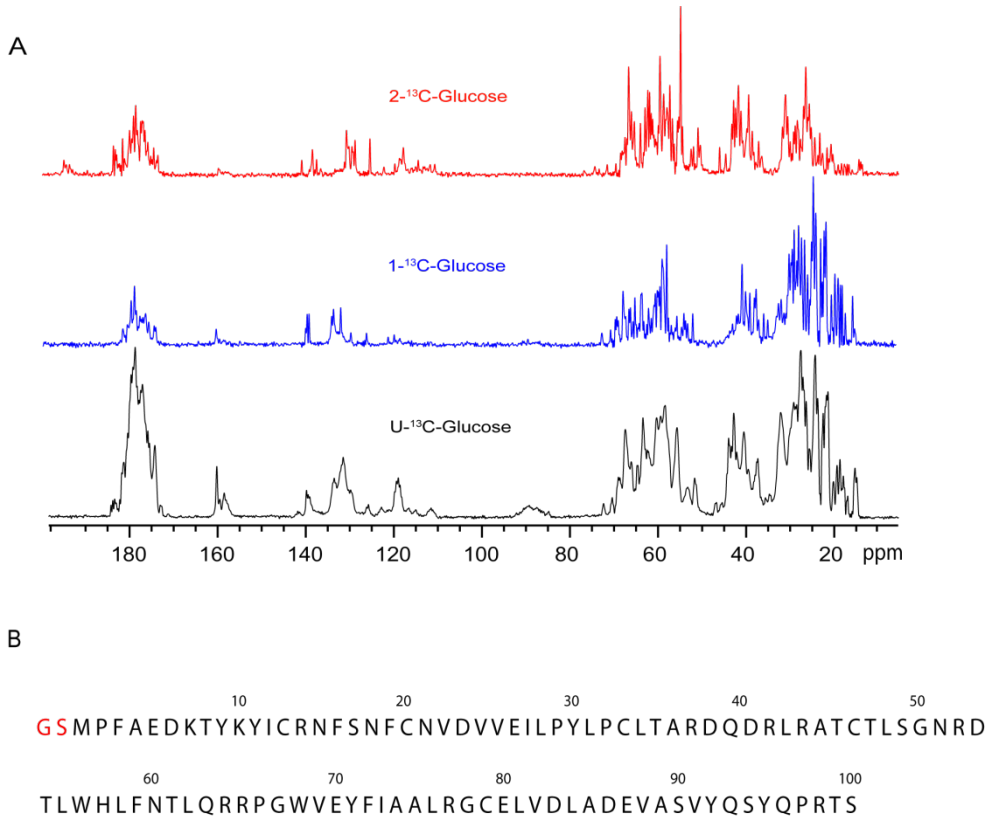

Figure 31. 1D ^{13}C spectra of MAVS-CARD$_{1\text{-}100}$ and its sequence: A. The spectra were recorded at a field of 600.28 MHz, a spinning frequency of 13.5 kHz and at 278 K, using a standard ^{1}H-^{13}C cross polarization (90–100% ramp on ^{13}C). **B.** The first two residue GS are the remaining residues after cleavage of the SUMO* tag.

3.3.2.2 Sequence specific resonance assignment

The sequential backbone assignments and most of the side chain assignments were obtained from a single uniformly labeled sample. Three 3D chemical shift correlation experiments (CONCA, NCACX , NCOCX, where CX denotes any carbon) and one 2D DARR experiment (mixing time 25ms) were acquired for this purpose. While the CONCA spectrum provided almost complete backbone resolution and gives the sequential information of CO[i-1], N[i], and C_{α}[i] atoms, the 2D DARR spectrum provided the 'finger print' pattern for most amino acid types. Glycines do not have side chains, however, they did provide unique cross peaks between C_{α} and CO. The NCACX and NCOCX spectra facilitated the sequential backbone walk and identified the side chain resonances correlated to their own backbone ^{15}N in NCACX and the following amino acid backbone ^{15}N in the NCOCX spectrum. These two 3D experiments alleviated overlap between side chain resonances by

virtue of the third dimension ^{15}N chemical shift. To illustrate sequential assignments, a representative strip plot of NCACO, NCACX and NCOCX is shown in **Figure 32**. In the 120.9 ppm strip of the

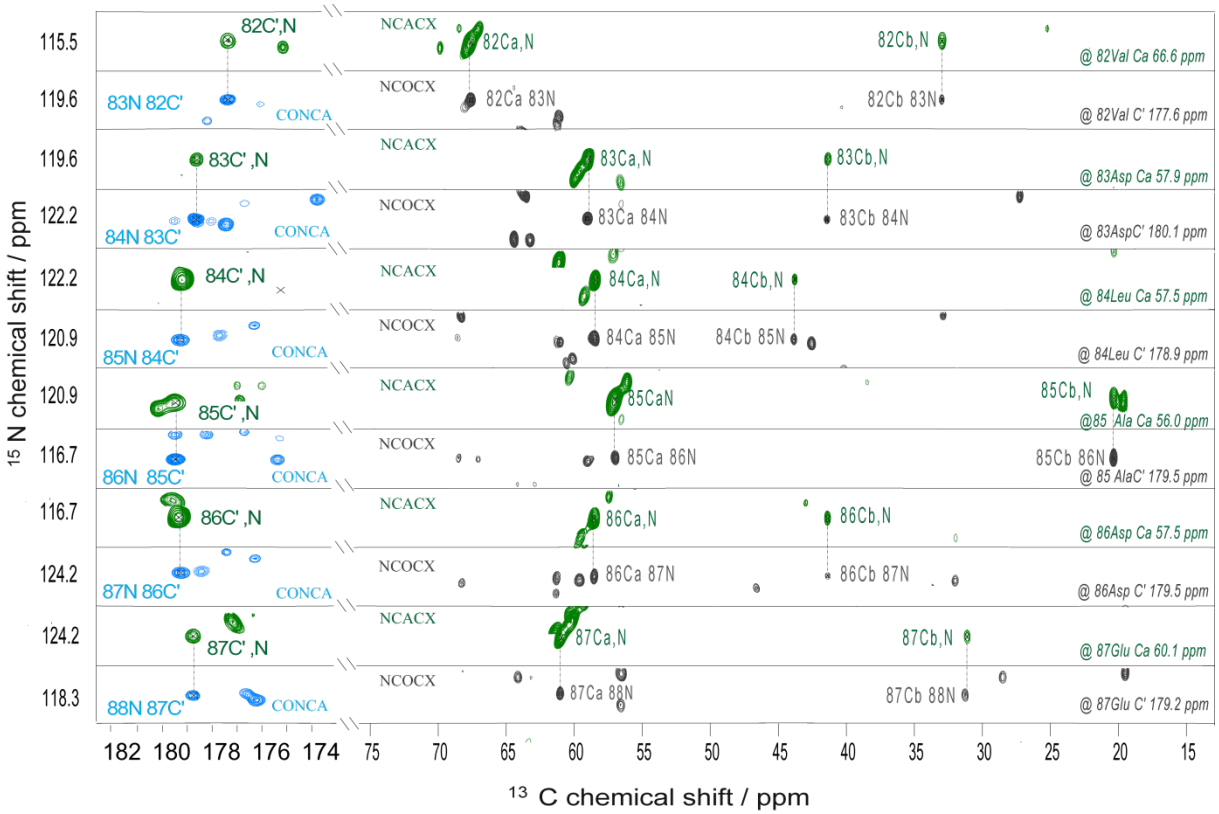

Figure 32 . An example of a sequential assignment walk for residues 82 to 87 of MAVS-CARD: Strips plot extracted from 3D CONCA, NCACX and NCOCX experiments illustrated the sequential assignment strategy. The NCACX spectrum is shown in green. The NOCX spectrum is shown in black and the CONCA spectrum is shown in cyan.

NCACX experiment (**strip 7 in Figure 32**), which contained information about N[i], C_α[i] and CO[i], one alanine spin systems could be identified. To determine which of the seven alanine residues in the sequence this spin system corresponded to, the 3D CONCA spectrum was then used to search for CO[i-1] chemical shift from N[i], and Ca[i] chemical shifts. This allowed to identiy the CO[i-1] chemical shift at 178.9 ppm. In the next step, based on the N[i] and CO[i-1] information, the 3D NCOCX experiment which contained information about N[i], CO[i-1] and CX [i-1] was used to identify the previous amino acid type (CX [i-1]) from the cross peak pattern which turned out to be a leucine. In the sequence, 84 leucine was the only leucine followed by alanine, thus the connections and assignments could be made. Using such a strategy, sequential connection was established for each amino acid. For the aromatic amino acid side chains, many of their chemical shifts were

assigned directly from the 2D PDSD spectra of 1-^{13}C glucose and 2-^{13}C glucose sparsely labeled sample due to the high resolution of the spectra (**Figure 40**). To confirm this assignment, 300 ms NCACX and 250 ms NCOCX spectra were recorded and analyzed on uniformly labeled sample, which contained cross peaks between two sequential residues (**Figure 39**). In the long mixing time NCACX experiment, $^{13}C_{\alpha}$ not only exchanged with the side chain carbons of the same residue but also with the backbone and side-chain spins of the sequential previous residue. Therefore this provided a way to check the self-consistency of the sequential assignment. In complete analogy, for long mixing time NCOCX spectrum, the ^{13}CO [i-1] could also transfer to the side chain of the previous residue [i].

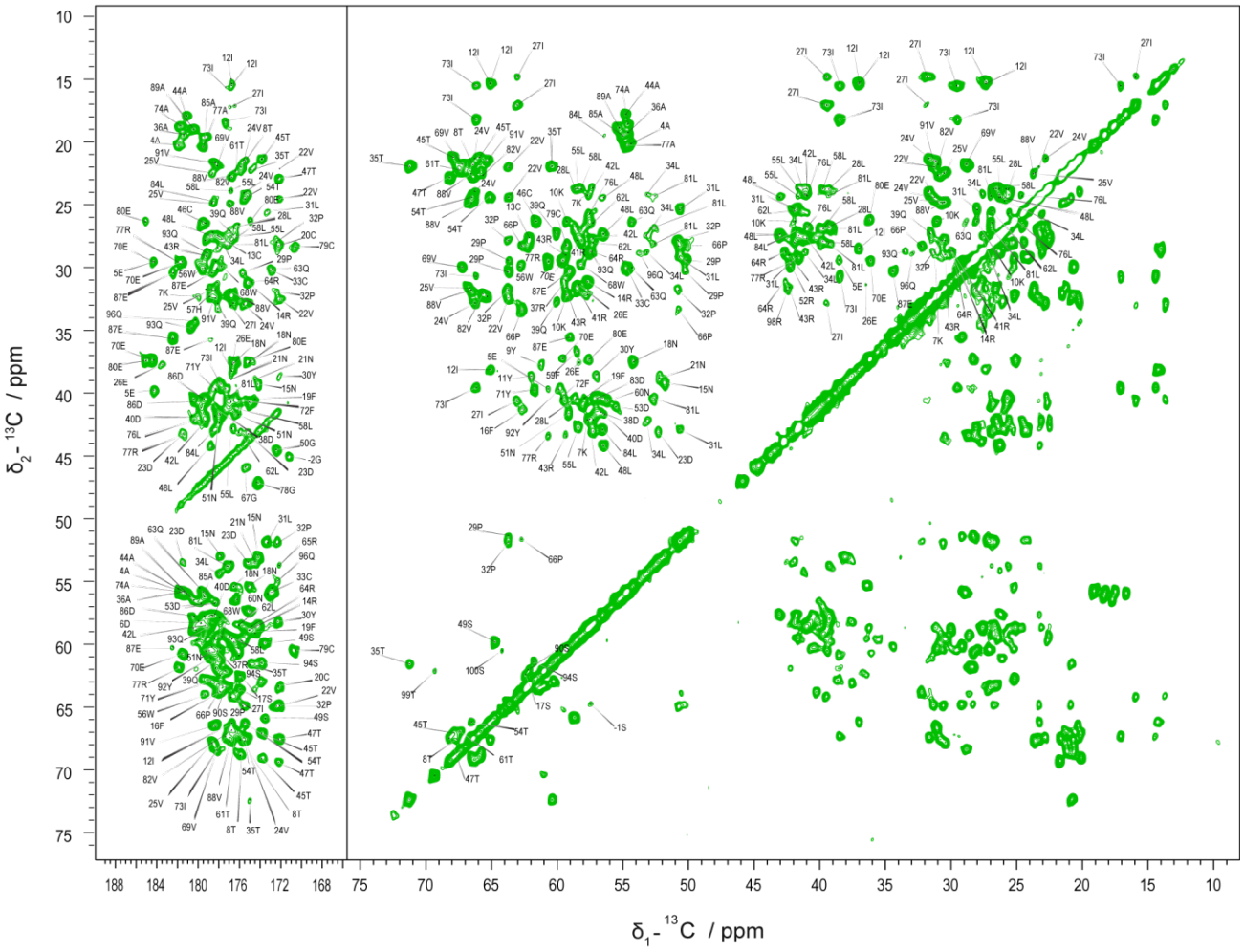

Figure 33. Contour plot of 2D ^{13}C-^{13}C DARR spectrum: 2D DARR experiment of MAVS-CARD with 25ms mixing time. The aliphatic region and carbonyl region are shown together with chemical shift assignments, which were achieved by analysis of data sets containing there dimensional heteronuclear NCACX, NCOCX and CONCA spectra.

Figure 33 shows a 25ms DARR ^{13}C-^{13}C spectrum on uniformly labeled sample(Takegoshi et al., 2001), which displayed only little spectral overlap in both the aliphatic and carbonyl regions. The entire set of cross peaks could be assigned by the strategy described above. Finally, from the aforementioned spectra 101 residues of MAVS-CARD$_{1-100}$ (containing GS in the N terminus) were assigned. 92.25% of all carbons were assigned. The unassigned carbons were from the N and C termini, which were flexible in solution (**Figure 23**). All resonances in the NCA spectrum could be assigned (**Figure 34**), corresponding to 92% of all backbone amide groups. Interestingly, double sets of C_β chemical shift for some aromatic amino acids, especially tyrosine, could be observed. The detected chemical shift differences between the double conformers were within the range of 0.2- 0.5 ppm (**Table 12**), which might indicate there were two slightly different conformations or slow exchange dynamics in the filamentous state.

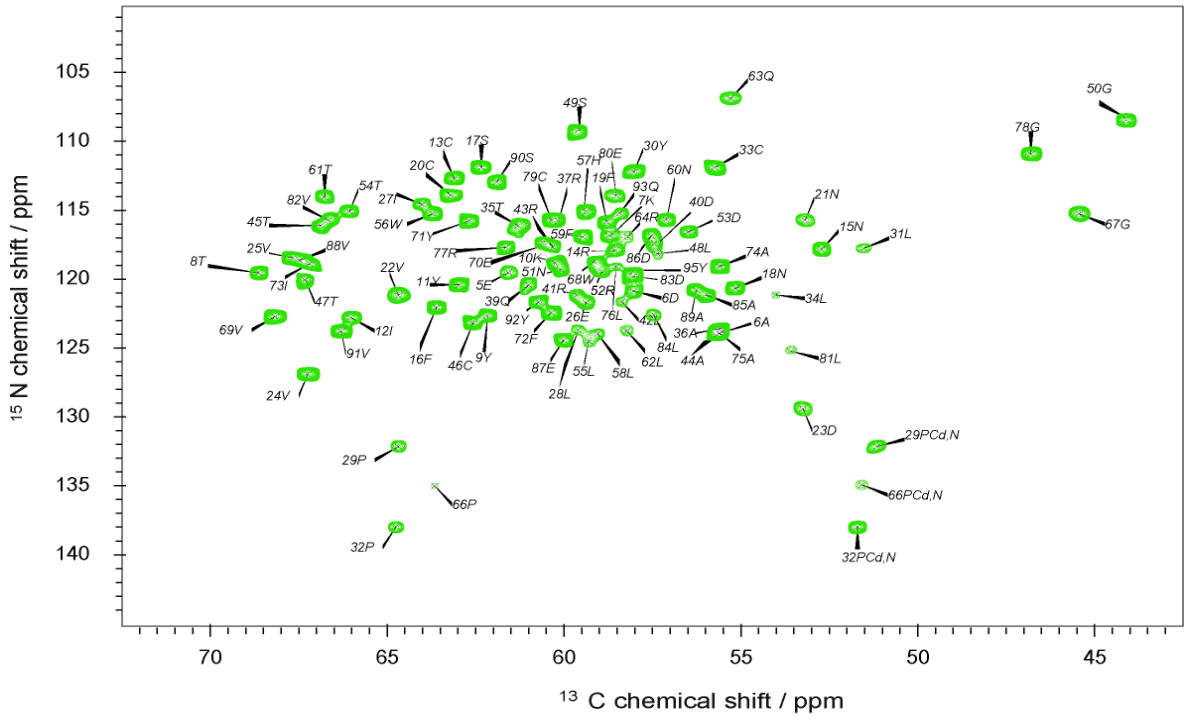

Figure 34. Contour plot of a 2D NCA spectrum of MAVS-CARD$_{1-100}$: The spectrum was acquired at 600 MHz proton frequency at a spin frequency of 13.5kHz and at a temperature of 278K. The data was analyzed with CCPN (Vranken et al., 2005). Assignment are also shown. For proline residues, the N, C_d cross peaks were also assigned.

Table 12. Double sets of C_β chemical shift of some TYR and PHE residues

AA	C_α (ppm)	C_β (ppm)	C_γ (ppm)
13 TYR	63.06	38.5/38.0	131.0
18 PHE	63.59	41.19/40.8	139.6
21 PHE	58.85	40.28/40.1	141.39
32 TYR	58.03	38.57/38.41	131.1
73 TYR	62.69	39.60/39.33	130.7

3.3.2.3 Secondary structure from the chemical shift index (CSI)

Based on the sequence of the protein, the backbone ^{13}C chemical shifts and TALOS+ (Shen et al., 2009) were used for experimental prediction of the secondary structure. This approach exploits the fact that the chemical shifts of the residues are affected by the local environment of the nuclei and correlate with the dihedral angles (Spera and Bax, 1991). Especially the chemical shift of the C_α and

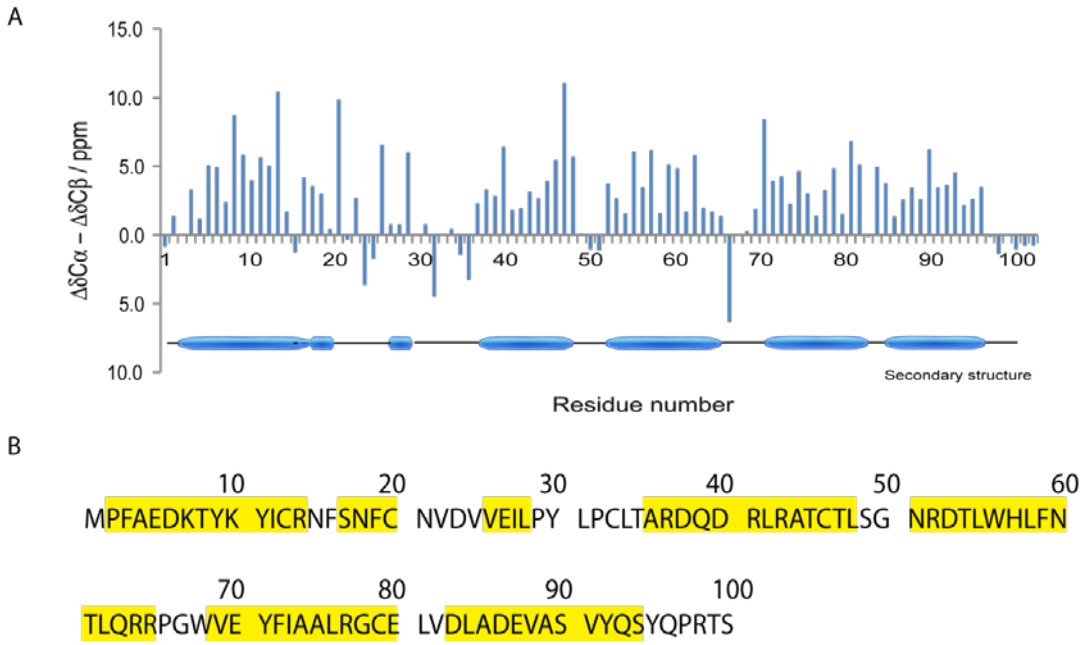

Figure 35. Chemical shift index (CSI) analysis: A. Secondary chemical shift difference plotted as a function of residue number. Three or more positive shifts in a row indicates a helix secondary structure. **B.** Sequence based representation of

the secondary structure of MAVS-CARD$_{1-100}$ predicted by TALOS+ (Shen et al., 2009), where the α-helices are heighted in yellow.

C_β are good indicators for the secondary structure of a protein. TALOS+ compares the secondary chemical shifts of a patch of residues from the query protein with proteins of known structure in the Protein Data Bank (PDB) database to determine the torsion angles and secondary structure. Based on the chemical shifts of MAVS CARD, 6 helices were predicted, comprising residues 2-14 (helix 1a), 16-20 (helix 1b), 25-28 (helix 2), 36-48 (helix 3), 51 to 65 (helix 4), 69 -80 (helix 5) and 82-94 (helix 6), as shown in **Figure 35**. More than three consecutive positive values in a row indicate α-helical conformation, three or more negative ones indicate a β-strand (Wishart and Sykes, 1994). Most of the assigned residues were found in α-helical conformation as depicted by the representative schematic below the plot. In helix 1, the chemical shift deviation of Asn15 was negative, which was consistent with the determined solution NMR structure which displayed a kink in helix 1 at residue 15. In summary, the CSI (Wishart and Sykes, 1994) indicated the existence of six helices separated by five short turns, in addition to the two terminal stretches. Thus, MAVS-CARD kept the same secondary structure in the filament as in its inactive monomeric state. This is in contrast to the prion protein PrP, that undergoes a significant change in secondary structure from α helix to β sheet to form fibrillar amyloids.

3.3.2.4 Chemical shift perturbation between monomer and filament state MAVS-CARD$_{1-100}$

The chemical shifts of filamentous MAVS-CARD$_{1-100}$ were also compared to the corresponding chemical shifts of monomeric MAVS mutant D25S E28S, which were recorded in the same buffer. As the chemical shift is determined by the local environment, the residues at the interaction interfaces of MAVS-CARD protomers in the filament would result in a different chemical shift. Since the peak width for solid state NMR spectra was broader, chemical shift tolerance was usually kept as 0.2 ppm. Perturbations of the C_α and C_β chemical shifts larger than ± 0.4 ppm (double tolerance) in the filamentous state relative to the monomeric state were considered to be significant. Base on these criteria, a series of residues were identified as highly affected by the formation of the filament (**Figure 36**). The surface plot of these residues on the solution NMR structure of MAVS-CARD$_{1-100}$ revealed they were located in parts of helix1, helix 2, helix 2- helix 3 loop, helix 3, helix

3- helix 4 loop and part of helix 4 (**Figure 36**). These regions are thus expected to be at the interaction interfaces in the MAVS-CARD filament.

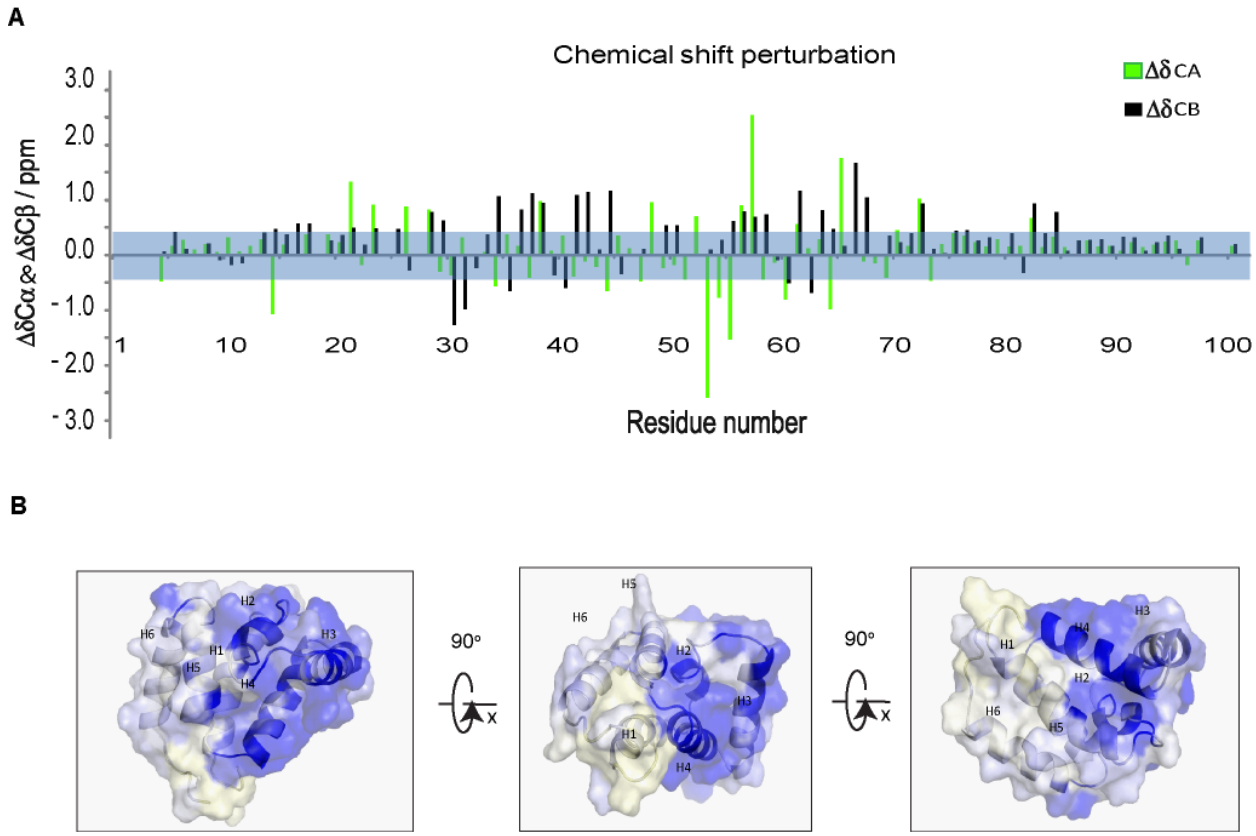

Figure 36. Chemical shift perturbation and surface plot of the perturbed residues: A. the chemical shift perturbation between monomeric state mutant D25S E28S and fibrillar state MAVS-CARD$_{1\text{-}100}$. **B.** The residues that had significant difference of either C_α or C_β chemical shift between two states of MAVS CARD are plotted in blue.

3.3.3 Structural calculation of MAVS-CARD$_{1\text{-}100}$ within the filament

In the **section 3.3.2**, the strategies of achieving unambiguous assignment of ^{15}N, ^{13}C resonances and determination of secondary structure have been described. This section will present the way for getting distance constraints and calculation of the tertiary structure of fibrillar MAVS-CARD using ARIA (Bardiaux et al., 2008; Rieping et al., 2007). The quality of the calculated structure relies on the accuracy and number of the distance constraints, especially medium and long-range distance constraints. For this purpose, the prior knowledge about the secondary structure will be helpful for

determining short and medium or long-range distance constraints. The accuracy of the constraints depends on the quality of the resonance assignments, as false assignments would be misleading during the structure calculation. However, a limited number of ambiguous or even wrong assignments can be handled by the structure calculation algorithm, giving rise to violations that can be reassigned during the refinement process. A scheme for structure calculation from solid state NMR data is showed in **Figure 37**.

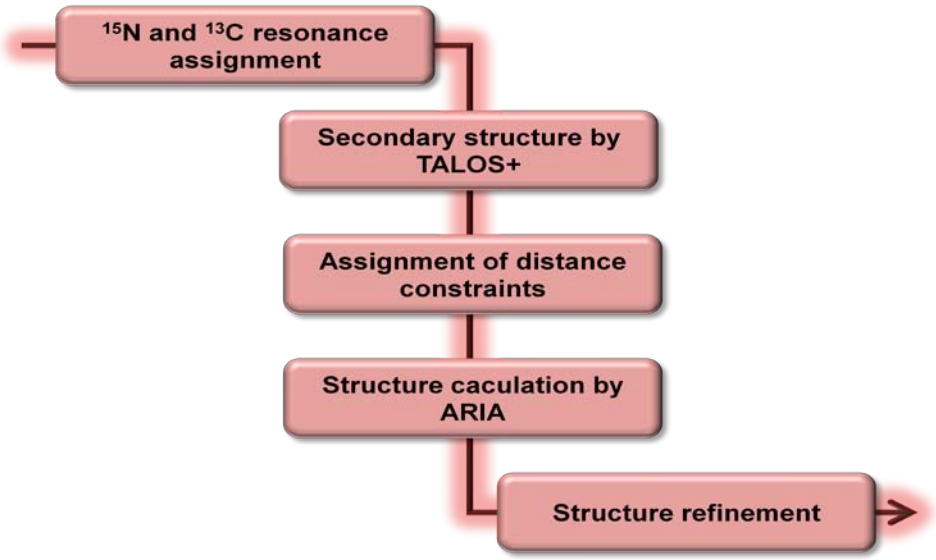

Figure 37. Schematic representation of steps in calculation of a protein structure by solid state NMR: ARIA could use both unambiguous and ambiguous distance constraints together with dihedral angles derived from TALOS+ for tertiary structure calculation.

3.3.3.1 Distance constraints from uniformly labeled sample

Distance constraints between side-chain or backbone atoms were required for determination of the tertiary structure of MAVS-CARD$_{1\text{-}100}$ in the fibrillar form. To get enough distance constraints, a set of homonuclear and heteronuclear spectra with different mixing times were recorded on a uniformly labeled sample and on sparsely labeled samples. 2D ^{13}C-^{13}C DARR or PDSD spectra provided a lot of structural information. A series of such spectra with increasing mixing times revealed intra-residue, sequential residue, medium and long distance contacts between carbon atoms. Principally, the higher the mixing time, the more and longer distance will be observed. In 50 ms, 100 ms and 150

ms DARR spectrum, intra-residue and sequential residue constraints were mainly observed. With longer mixing times, i.e. 250 ms and 400 ms DARR spectra, a lot of medium and long-range distance constraints could be observed (**Figure 38**), whereas many intra-residual peaks disappeared due to relaxation processes.

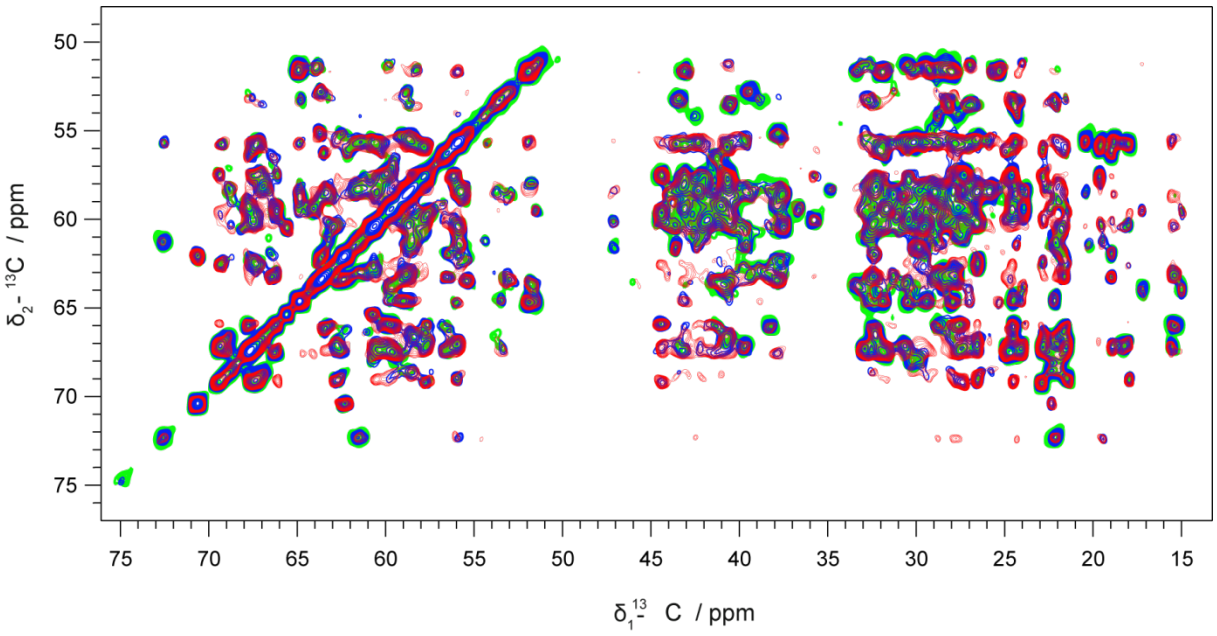

Figure 38. Peak information of DARR spectra depends on the mixing time: Superposition of 2D DARR spectra with 50 ms (green), 150 ms (blue), 400 ms (red) mixing time. Many long-range constraints showed up only in 400ms spectrum.

The largest number of constraints could be obtained from carbon-carbon correlation spectra of a uniformly ^{15}N and ^{13}C labeled sample. However, unambiguous assignment of these constraints would be difficult due to the overlap of the peaks. Only a few peaks were unambiguous in both dimensions. Some peaks were unambiguous in one dimension and the other dimension was ambiguous. To reduce the ambiguity, a third dimension was introduced by acquiring 250ms/300ms NCOCX and NCACX spectra, separating the overlapping peaks via the ^{15}N chemical shift (**Figure 39 A**). In order to avoid the risk of mis-assigning short-range constraints as long-range constraints, peak intensities in different mixing time DARR spectra were taken into consideration. Usually the short-range constraints were observed in lower mixing time spectra, while the long-range constraints were only observed in longer mixing time spectra. In addition, intra-residue and sequential residue assignment were given preference over the medium and long-range assignments. Since the secondary

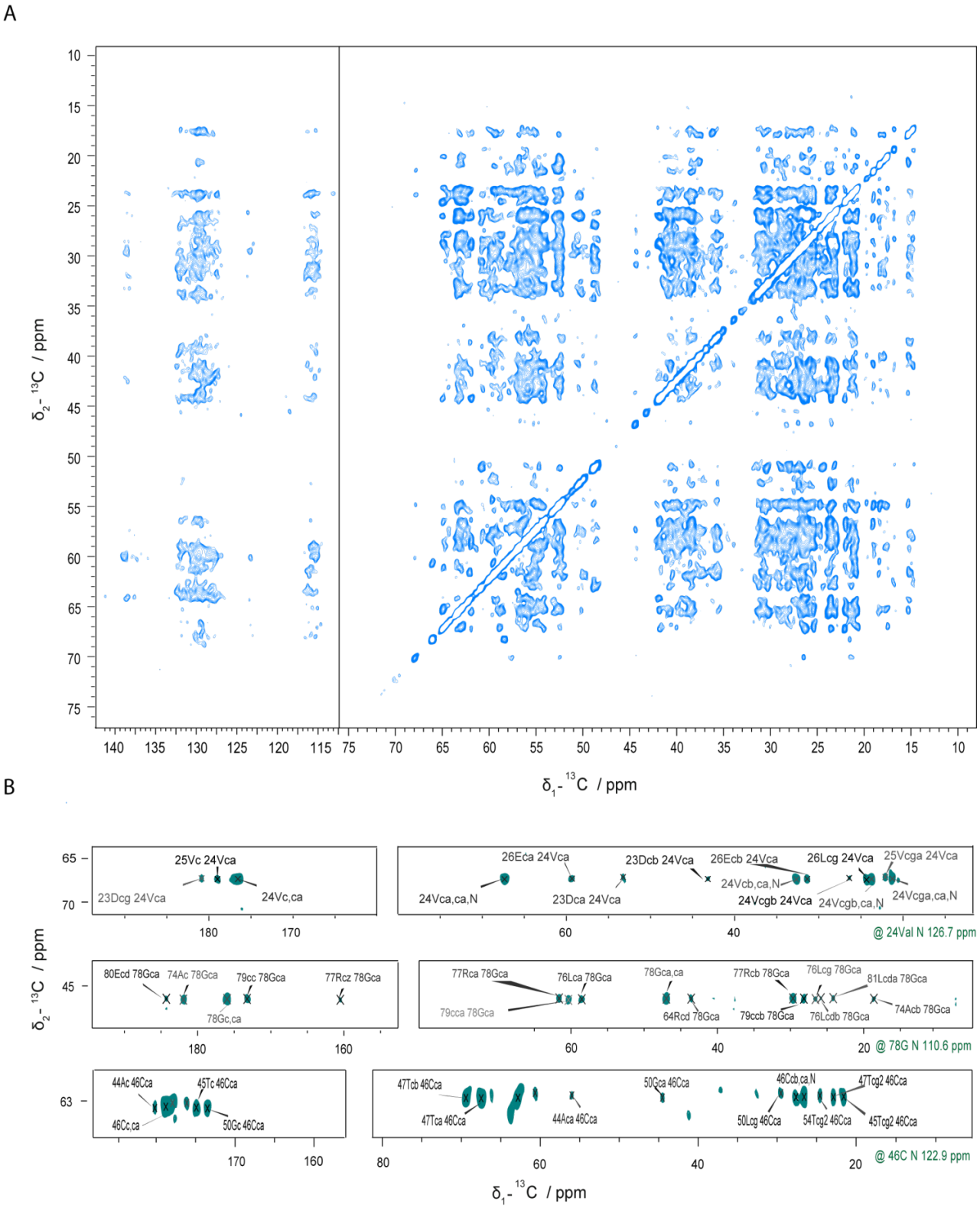

Figure 39. Contour plots of 2D DARR spectrum and 3D 300ms NCACX spectrum: A. 2D DARR spectrum of a 1:6 diluted sample with a mixing time of 400 ms. Due to the limited resolution of the spectrum, most of the peaks were highly ambiguous. **B.** 3D NCACX spectrum with assignment. Introducing ^{15}N as the third dimension helped resolve the peaks that overlapped in 2D spectra. Several sequential, short and medium range assignments were made unambiguously.

structure of MAVS-CARD was mainly alpha helical, the constraints between residues "i" and "i ± 3", which defined the secondary structure of alpha helices, were observed quite often. Finally, to collect only intra-molecular constraints for structure calculation of a MAVS-CARD protomer, a diluted sample at a molar ratio of 1:6 was made (**section 2.2.2**). 822 intra-molecular constraints were identified from this diluted sample (**Figure 39 B**).

3.3.3.2 Distance constraints from sparsely labeled sample

In addition to three-dimensional (3D) spectroscopic techniques, the ^{13}C spin dilution strategy also provided an additional tool to alleviate the spectral overlap. In this project, the ^{13}C diluted protein samples were produced using 1-^{13}C and 2-^{13}C glucose as the sole carbon source in the medium (Loquet et al., 2011; Lundstrom et al., 2007). This labeling schemes reduced the carbon-carbon dipolar coupling and resulted in favourable ^{13}C spectral resolution (**Figure 31**). This facilitated the assignment and collection of distance constraints as previously reported glycerol based labeling schemes (Hong, 1999; LeMaster, 1996). Notably, the 1-^{13}C glucose or 2-^{13}C glucose based labeling pattern was more sparse than that obtained by using 2-^{13}C glycerol or 1,3-^{13} C glycerol as the sole carbon source, because only one of the six carbons in glucose was ^{13}C labeled. Numerous cross-peaks could readily be assigned to medium and long-range distance constraints. The two labeling schemes were almost complementary to each other. The spectra of 1-^{13}C-glucose labeled sample contained many methyl-based distance constraints. Similarly, the 2-^{13}C-glucose sample provided a lot of C_α based constraints and showed excellent resolution. The high quality of MAVS-CARD spectra allowed a precise detection of the resonance frequencies (±0.1ppm), permitting the collection of unambiguous constraints in both aliphatic and aromatic regions (**Figure 40**). In total, 620 peaks were assigned unambiguously (including 46 long-range constraints) from ^{13}C-^{13}C correlation spectra of 1-^{13}C and 2-^{13}C glucose labeled sample. The remaining 456 peaks were kept ambiguous either in one dimension or in two dimensions. The unambiguous constraints between F19 with I12 confirmed the kink segment that observed from secondary chemical shift (**Figure 40**).

After the acquisition of ^{13}C-^{13}C correlation spectra, the application of 1-^{13}C and 2-^{13}C glucose was also extended by using of the ^{15}N dimension. A 2D PAIN-CP spectrum was acquired in a relatively short time (40h) on the 1-^{13}C glucose labeled sample, showing excellent resolution in the ^{15}N-^{13}C dimension (**Figure 41 A**). For example, two close resonances, 74A C_β (18.53ppm) and 89A C_β (18.67ppm) were separated from each other quite well (**Fig 41 B**). Many sequential, short, medium

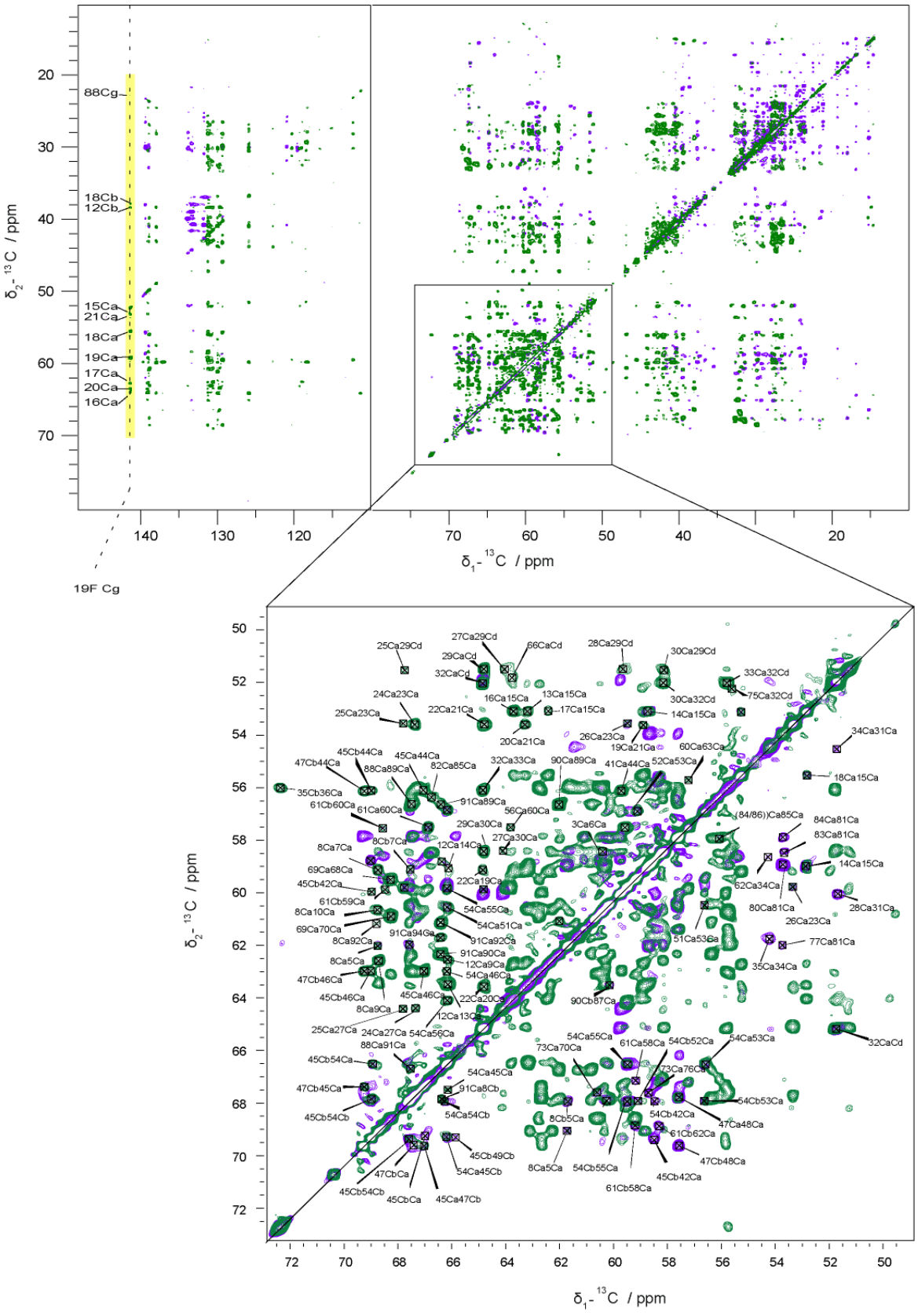

Figure 40. 2D PDSD spectra of 1-^{13}C and 2-^{13}C glucose labeled sample with a mixing time of 1000 ms: The spectrum in green was recorded on a 2-^{13}C glucose labeled sample and the spectrum in mauve on a 1-^{13}C glucose labeled

sample. The two labeling schemes were almost complementary to each other. The high resolution in both the aliphatic and the aromatic region facilitated the assignment of distance constraints. In the lower part, the Cα region from DARR spectra of both 1-^{13}C and 2-^{13}C glucose labeled sample are shown with assignment. The unlabeled peaks were used as ambiguous constraints.

and long- range constraints could be assigned unambiguously. In the end, 256 peaks out of 540 peaks were assigned, out of which 18 peaks were long-range constraints.

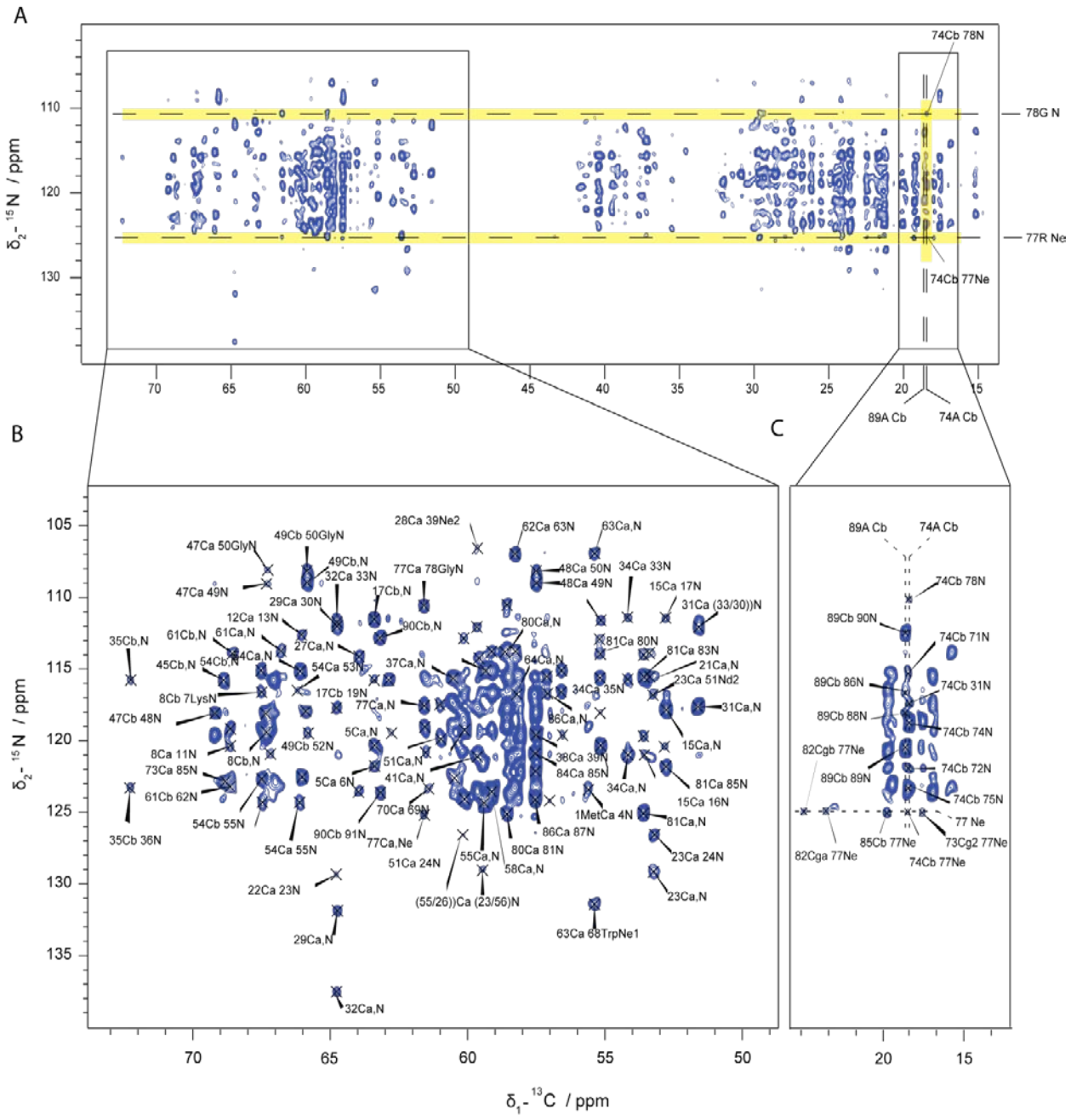

Figure 41. 2D PAIN spectra of 1-^{13}C glucose labeled sample: A. Overall view of the PAIN spectrum showed highly resolved peaks. **B.** The NC_{α} region is enlarged and shown together with assignments. The unlabeled peaks were taken as

ambiguous constraints. **C.** The high resolution of the spectrum allowed a precise detection of the resonance frequencies (±0.1ppm). 74A C_β and 89A C_β had chemical shifts of 18.53ppm and 18.67ppm respectively.

3.3.4 *De novo* structure calculation of MAVS-CARD$_{1-100}$ in fibrillar state by ARIA

The structure of the MAVS-CARD$_{1-100}$ protomer within its fibrillar assembly was determined using ARIA with the assistance of Dr. Benjamin Bardiaux, who developed custom-made extensions to the software. ARIA was originally written for automatic NOE assignment and NMR structure calculation. The structure determination of MAVS-CARD$_{1-100}$ within its filamentous assembly was challenging because of several reasons. Firstly, the MAS solid state NMR distance constraints, in general, were highly ambiguous and less defined compared to typical solution NMR data. Secondly, the fibrils were usually less ordered than microcrystalline samples previously studied by solid state NMR. Thirdly, because MAVS-CARD$_{1-100}$ assembled into a fibrillar structure, one needed to distinguish intra- and inter-molecular constraints by making aforementioned diluted and mixed samples (**section 2.2.2**) Finally, NMR-based information is intrinsically of local nature, with a distance limit of approximately 7-8 Angstrom. Thus, additional mesoscopic information is typically required, such as symmetry parameters for the fibrillar assembly, or EM-derived information on the fibrillar assembly. In this section, only the structure calculation of the MAVS-CARD$_{1-100}$ protomer in its filament assembly will be described. The quaternary structure of the MAVS-CARD$_{1-100}$ assembly will be described in **section 3.5.2**. Distance constraints were derived from spectra of 4 samples with different labeling schemes, including PDSD spectra of 1-^{13}C glucose and 2-^{13}C glucose labeled sample with mixing times of 1000 ms, DARR spectra of a uniformly ^{13}C,^{15}N labeled sample and of a diluted ^{13}C,^{15}N sample with mixing times of 400 ms, and a PAIN spectrum of a 1-^{13}C glucose labeled sample. All the ^{13}C-^{13}C correlation spectra were compared to the spectrum of the diluted sample to pick only intra-molecular constraints. 2687 cross peaks were picked, out of which 876 are assigned unambiguously including 64 long distance constraints. CCPN software was used to generate all the distance constraints with a carbon chemical shift tolerance of 0.2 ppm and nitrogen chemical shift tolerance of 0.4 ppm for sparsely labeled samples. For the uniformly labeled sample, due to the lower resolution, the carbon chemical shift tolerance was set to 0.35 ppm and nitrogen chemical shift tolerance was set to 0.6 ppm. The same constraints appearing in two or more spectra were merged. A list of 2021 non-redundant constraints was obtained by combining all the spectra. The minimum and maximum distance bounds for all constraints were 1.7 Å and 8 Å respectively.

With these NMR distance constraints, the structure of MAVS-CARD monomer within its filament assembly was calculated by ARIA using the conventional iterative protocol (Rieping et al., 2007).

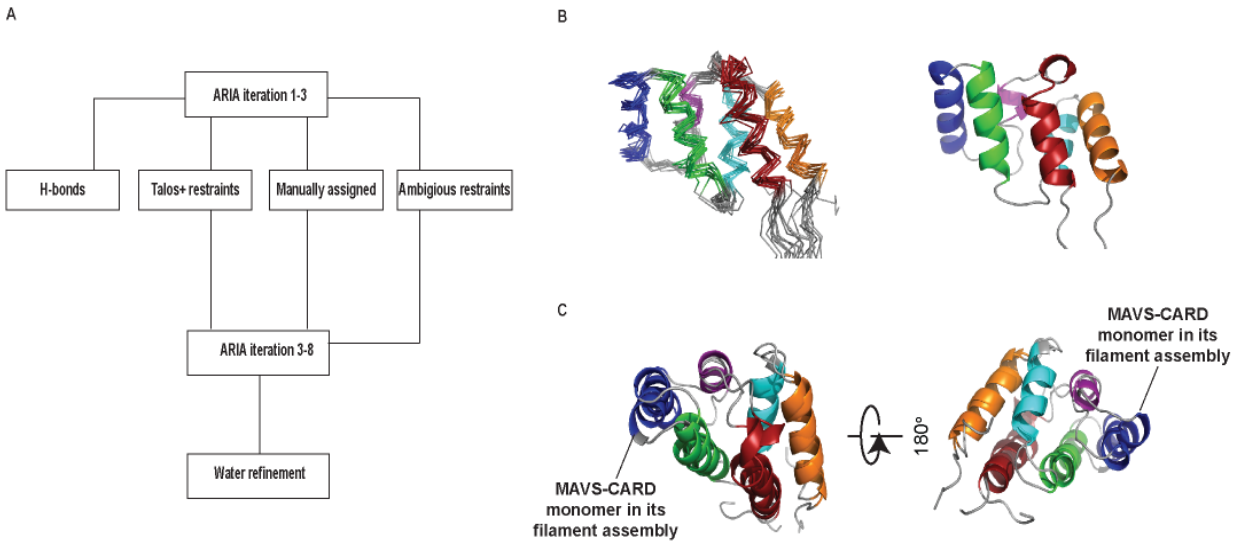

Figure 42. Structure of MAVS-CARD$_{1-100}$ in the fibrillar state calculated by ARIA: A. A representative scheme for using ARIA to calculate of the structure of MAVS-CARD monomer within its filament. H-bonds restraints were used in the first three iterations. **B.** MAS solid state NMR structure of MAVS-CARD$_{1-100}$ determined by ARIA. It forms a bundle of six helices in Greek key topology **C.** Overlay of the structure of MAVS- CARD$_{1-100}$ in monomeric state and in fibrillar state. The helix3 was pulled a bit outwards.

The general scheme of the structure calculation and the structure obtained are shown in **Figure 42**. All manually assigned unambiguous distances constraints, the remaining ambiguous distance constraints, dihedral angle restraints and H-bonds from TALOS+ were used to calculate the structure of MAVS-CARD$_{1-100}$ in its fibrillar state. The hydrogen bonds were used in the first three iterations to get a better defined secondary structure. The final ensemble of the 20 best structures showed high precision with RMSD of 0.99 ± 0.24 Å and 1.54 ± 0.29 Å for backbone atoms and all heavy atoms respectively. The statistics for the experimental constraints and distance constraints on each residues were shown in **Table 13** and **Figure 43** respectively.

Table 13. Summary of structural constraints and structure statistics.

Number of constraints
ssNMR distance constraints

Unambiguous	1298
Intra-residue (|i-j| = 0)	430
Sequential (|i-j| = 1)	418
Medium-range (2 ≤ |i-j| < 5)	298
Long-range (|i-j| ≥ 5)	152
Ambiguous	723
Total	2021
Dihedral angle restraints (φ/ψ)	182 (91/91)
Hydrogen bonds restraints	86
Constraints statistics[a]	
RMS of distance violations	
ssNMR constraints	0.011 ± 0.002 Å
H-bonds restraints	0.030 ± 0.004 Å
RMS of dihedral violations	1.074 ± 0.273 °
RMS from idealized covalent geometry	
bonds	0.003 ± 0.001 Å
angles	0.473 ± 0.017 °
impropers	1.157 ± 0.058 °
Structural quality[a]	
Ramachandran statistics[b]	
Most favoured regions	89.3 ± 1.7 %
Allowed regions	9.5 ± 1.8 %
Generously allowed regions	0.9 ± 0.6 %
Disallowed regions	0.4 ± 0.7 %
WHAT-IF Z-score[c]	
Backbone conformation	-1.32 ± 0.46
2nd generation packing quality	-1.82 ± 0.29
Ramachandran plot appearance	-1.94 ± 0.48
χ1/χ2 rotamer normality	-3.11 ± 0.46
Coordinates precision[d]	
All backbone atoms	1.61 ± 0.51 Å
All heavy atoms	2.09 ± 0.47 Å
All backbone atoms (5-96)	0.99 ± 0.24 Å
All heavy atoms (5-96)	1.54 ± 0.29 Å
Backbone RMSD vs. reference structure[e]	
Solution NMR	2.19 Å
X-ray (2VGQ)	1.97 Å

[a] Average values and standard deviations over the 15 conformers

[b] Percentage of residues in the Ramachandran plot regions determined by PROCHECK (Laskowski 1993)
[c] Z-scores values reported by WHAT-IF (Vriend, 1990)
[d] Average root mean square deviation (RMSD) over the 15 conformers' atomic coordinates with respect to the average structure.
[e] Root mean square deviation (RMSD) of the average ssNMR structure with respect to the reference structure for residues 5 to 95.

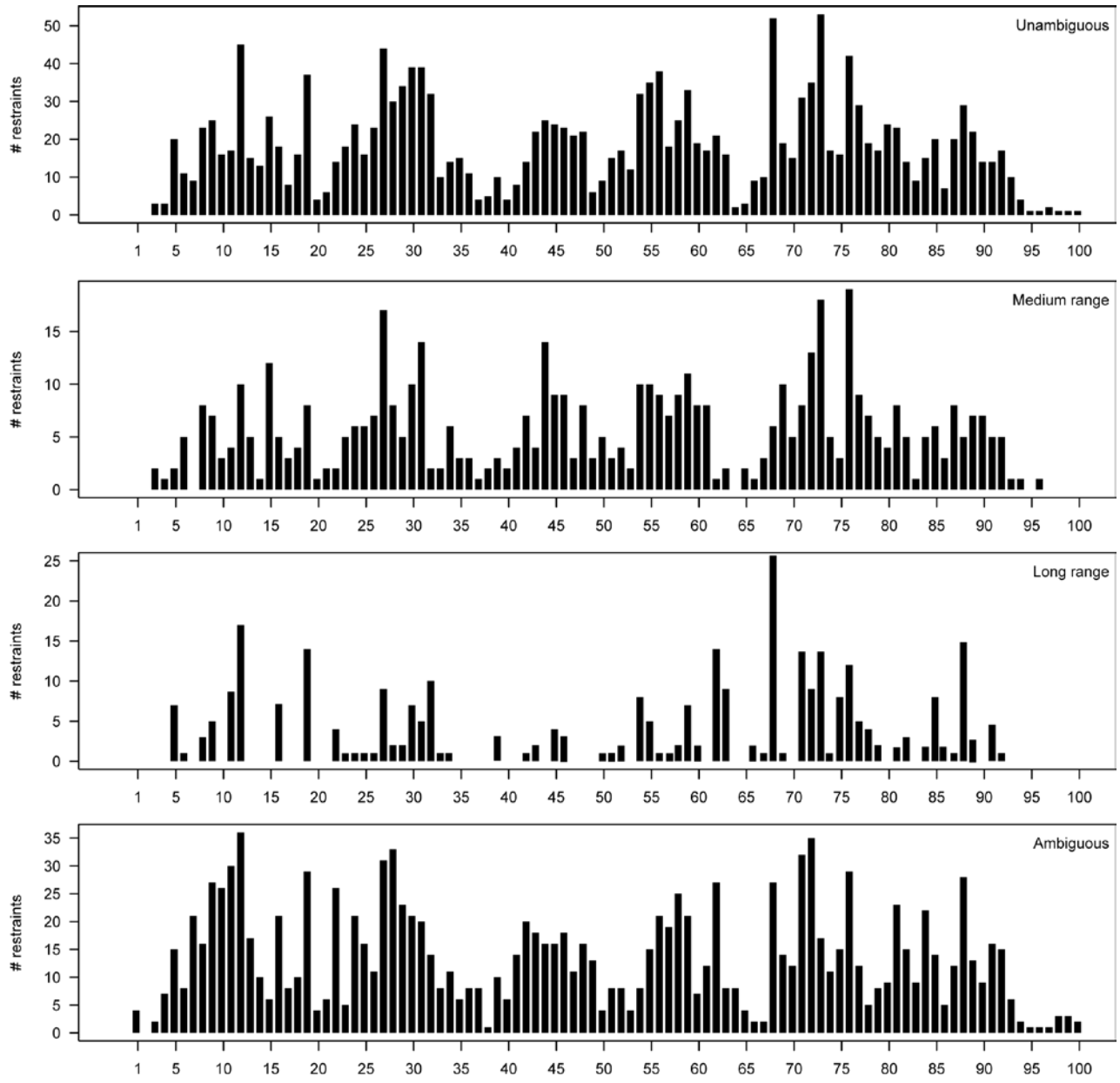

Figure 43. Statistics on distance constraints for each residue of MAVS CARD after ARIA structure calculation: The structure was calculated with manually assigned unambiguous distances constraints and the remaining unassigned ambiguous distance constraints. ARIA could assign the ambiguous peaks automatically. After structure calculation 1298

distance constraints were assigned unambiguously and 723 distance constraints were assigned ambiguously. The number of unambiguous medium and long-range distance constraints are plotted against the sequence.

The structure of MAVS-CARD$_{1-100}$ in its filamentous state was quite similar to that in its monomeric state, excepted that helix3 was pulled a bit outwards. In this respect, the activation of monomeric MAVS into its aggregated form was indeed different from the aggregation process found for the prion protein PrP, where, a transition from an α helix-rich monomer into a β sheet-rich filament was observed (Pan et al., 1993).

3.4 Assembly of MAVS-CARD$_{1-100}$ filament

3.4.1 Intermolecular distance constraints and three types of interaction

To distinguish the intra-molecular and inter-molecular constraints, a diluted sample (1:6), a mixed ^{15}N and uniformly ^{13}C labeled sample (1:1), and a mixed 1/2-^{13}C glucose labeled sample (1:1) were made (**section 2.2.2**). By comparison of the spectrum of the uniformly labeled sample with that of the diluted sample, 56 inter-molecular constraints were found, 10 out of them could be assigned unambiguously. In a similar way, 45 out of 76 intermolecular constraints were assigned from the 1-^{13}C glucose labeled sample and the 2-^{13}C glucose labeled sample. Overlay of the spectra of the 1-^{13}C glucose labeled sample, the 2-^{13}C glucose labeled sample and that of the mixed 1/2-^{13}C glucose labeled sample, 9 extra peaks from the spectra of the mixed sample were identified and assigned as inter-molecular constraints (**Figure 44**). Additionally, 2 more intermolecular constraints were identified from the ^{15}N and ^{13}C mixed sample. For the PAIN spectrum of 1-^{13}C-glucose labeled sample, the scheme, proton assisted insensitive nuclei cross polarization (PAIN-CP), reduced dipolar truncation and therefore was particularly suited for obtaining long distance constraints. The rejected peaks of this spectrum from ARIA for MAVS-CARD$_{1-100}$ monomer structure calculation were analyzed. 9 peaks were identified and assigned unambiguously as intermolecular constraints.

From all these intermolecular constraints, three different interaction interfaces could be identified, as it has been reported earlier that DDs use three mechanism of interactions (Ferrao and Wu, 2012; Park et al., 2007a). These interaction interfaces were named type I, II and III. The same convention was followed in the description of the interactions in this project (**Figure 45**). In type I interaction, residues at helix 1 and helix 4 interacted with residues at helix 2, helix 3 and the loop in between.

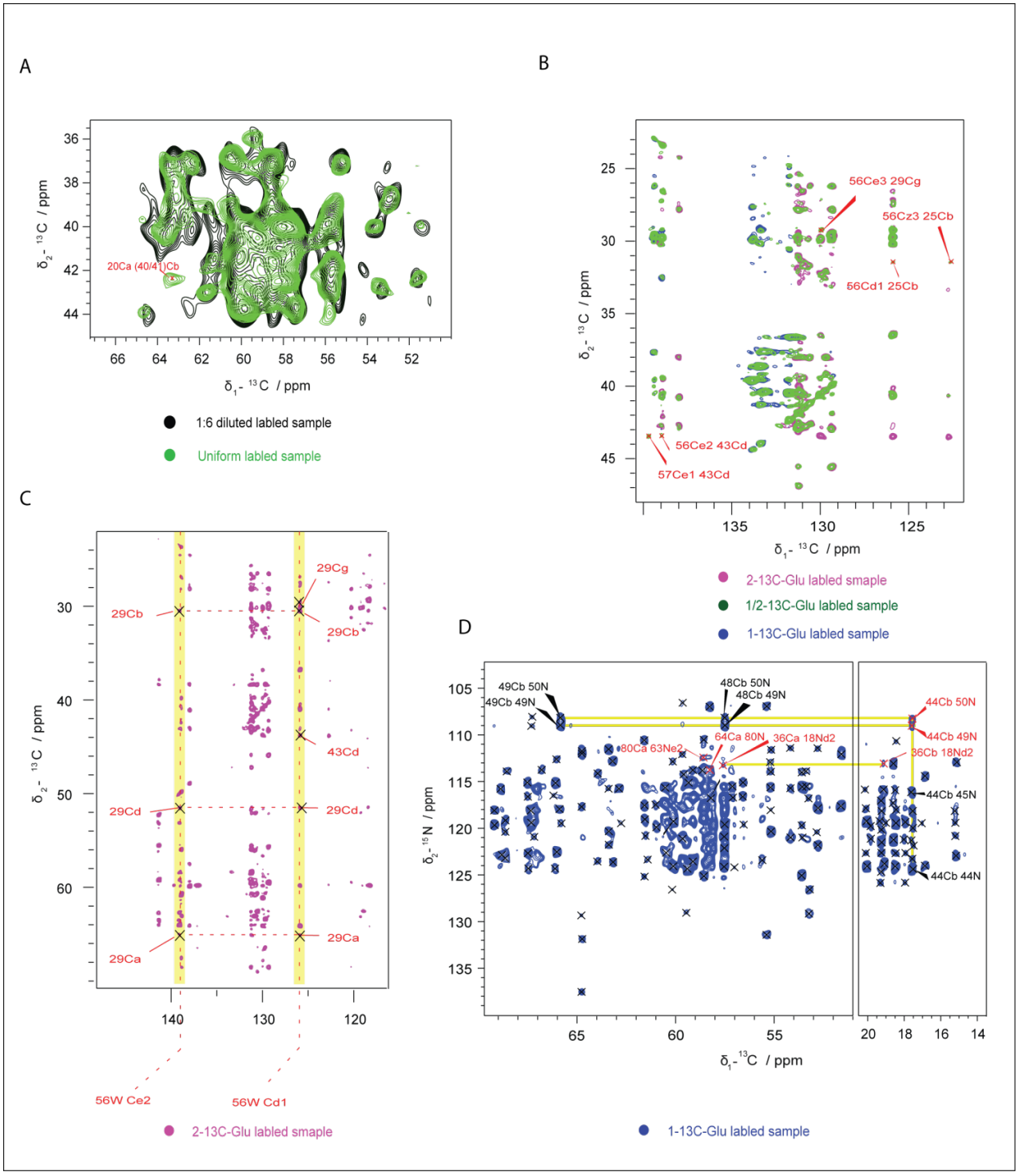

Figure 44. Intermolecular constraints were identified and assigned from different samples: A. The extra peaks in the uniformly labeled sample compared to that of the diluted sample were identified and assigned as intermolecular constraints. **B.** Overlay of the spectra from samples labeled by 1-^{13}C-glucose, 2-^{13}C-glucose and mixed 1/2-^{13}C-glucose helped to distinguish and assign intermolecular constraints. **C.** Inter-molecular constraints assigned on the spectrum from 2-^{13}C-glucose labeled sample. **D.** Inter-molecular constraints(red) assigned on the PAIN spectrum from 1-^{13}C-glucose labeled sample.

Many hydrophobic and polar interactions were involved in this interaction interface. Constraints for type I interaction between residues W56 and P29, R43 are shown in **Figure 44 B and C**. In total, 29 unambiguous and 10 ambiguous constraints from 14 different residues were found to support this interface. In the type II interaction, 10 unambiguous and 3 ambiguous constraints were identified, The residues Q63, R64, R65 of the helix4-helix5 loop and residues C79, E80, L81 from the helix5-helix6 loop formed this interface. It was found to be mostly polar and charged. Constraints shown in **Figure 44 D** between Q63, R64 with E80 belonged to this type of interaction. In the type III interaction, residues around helix1-helix2 loop and helix3-helix4 loop from one protomer interacted with residues of helix3 from another CARD. 15 intermolecular constraints from 8 different residues were identified. A mixture of electrostatic, polar and hydrophobic interactions occurred at this interface, such as constraints C20-D40/R41, N18-A36, A44-S49 and A44-G50 as shown in **Figure 44 A and D**. These six surface patches correspond well to those regions where chemical shift changes between the monomeric and filamentous state of MAVS-CARD$_{1\text{-}100}$ were observed (**Figure 36)**, thus confirming the conclusion that these chemical shift changes are caused by interactions between protomers.

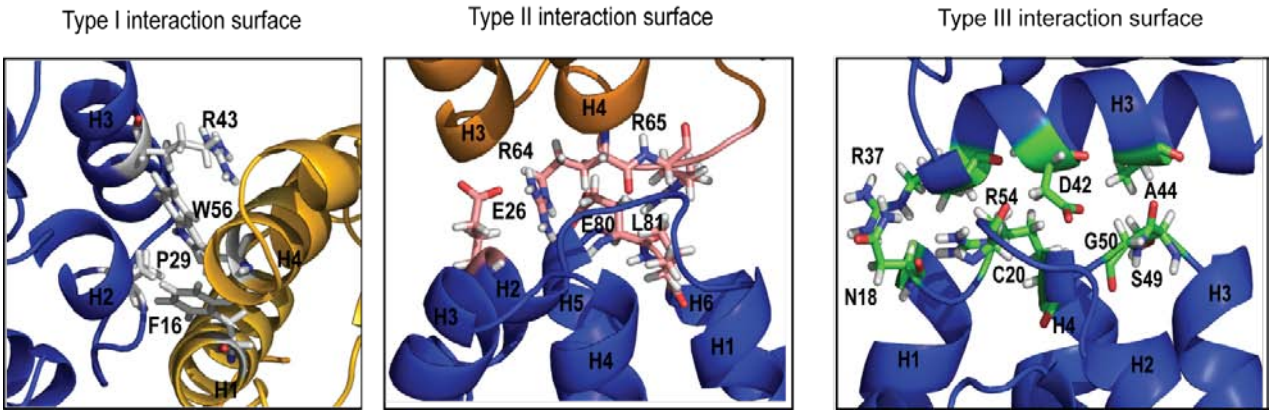

Figure 45. Three types of interaction surface involved in assembly of MAVS-CARD$_{1\text{-}100}$ filament: Important residues and their locations on MAVS-CARD$_{1\text{-}100}$ are labeled.

3.4.2 Probing surface accessibility of MAVS-CARD filament using paramagnetic relaxation

Solvent accessibility is typically analyzed in solution NMR for studying protein-protein interaction surfaces. In addition to mapping interaction surfaces by hydrogen-deuterium exchange,

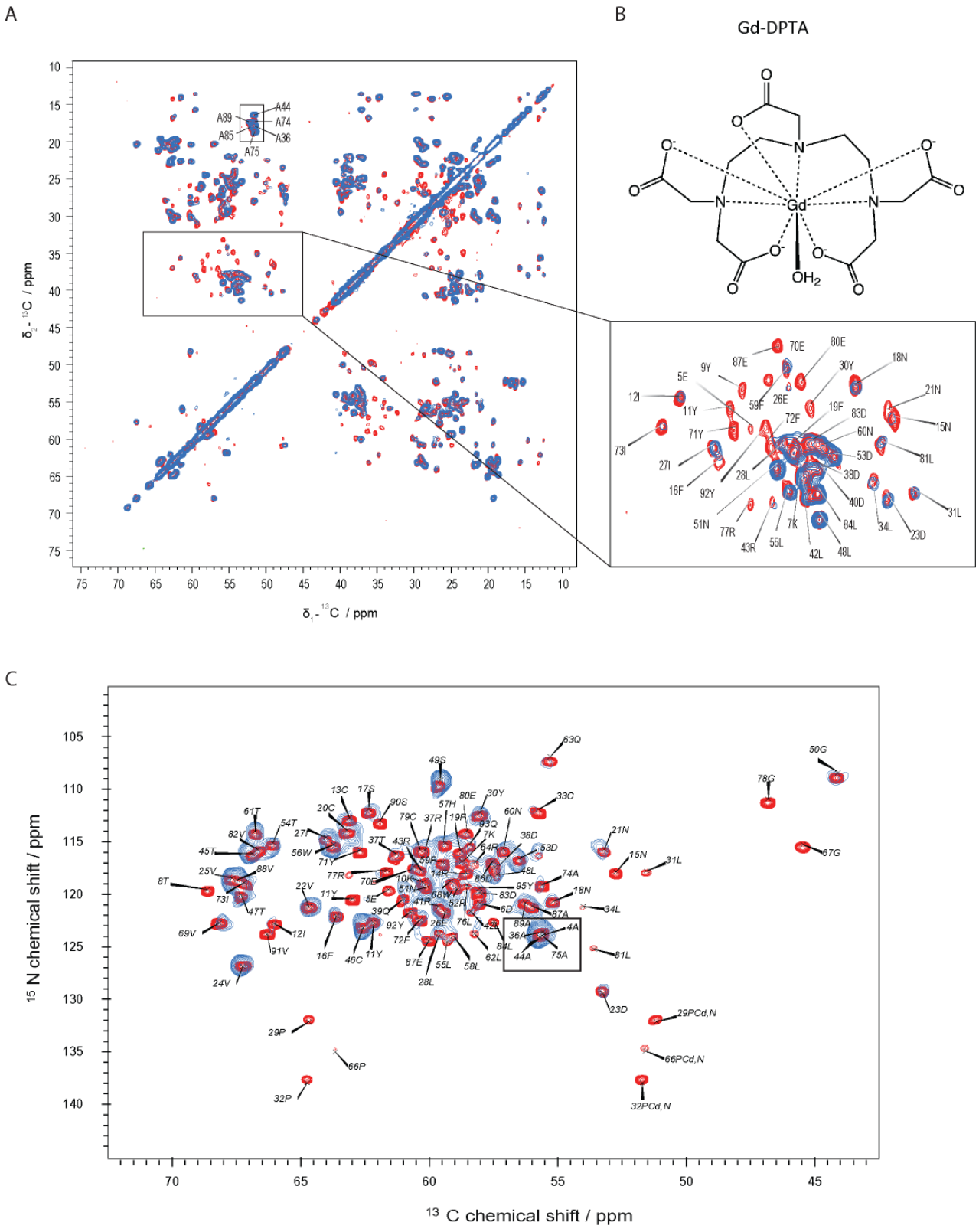

Figure 46. Surface accessibility probed by Gd-DPTA in solid state NMR: A. 25ms DARR spectra of MAVS-CARD$_{1\text{-}100}$ filaments in the absence (red) and presence (blue) of Gd-DPTA. The boxed region were enlarged and shown together with assignment. **B.** The molecule of Gd-DPTA. **C.** NCA spectra of MAVS-CARD filaments in the absence (red) and

presence (blue) of Gd-DPTA. The disappeared residues upon mixing with 100µM Gd-DPTA were located on the surface of the fibrillar assembly of MAVS-$CARD_{1-100}$.

(Paterson et. al., 1990), soluble radicals or paramagnetic ions could also be employed to map the solvent-exposed residues (Bernini et al., 2006; Linser et al., 2009; Ulmer et al., 2002). These cause a rapid relaxation of near-by spins in a strictly distance-dependant manner, thus affecting surface-exposed residues the strongest. The absence of specific interactions between paramagnetic probes and the proteins of interest is a prerequisite for this type of analysis. Gd-DTPA (**Figure 46**) as one of such probes was used to study the surface accessibility of proteins in this project. Due to its unpaired electrons, Gd-DTPA could bleach the resonance of the residues located on the surface of the protein. Thus, ^{13}C-^{13}C DARR spectra and ^{15}N-^{13}C NCA spectra of MAVS-$CARD_{1-100}$ filament were recorded

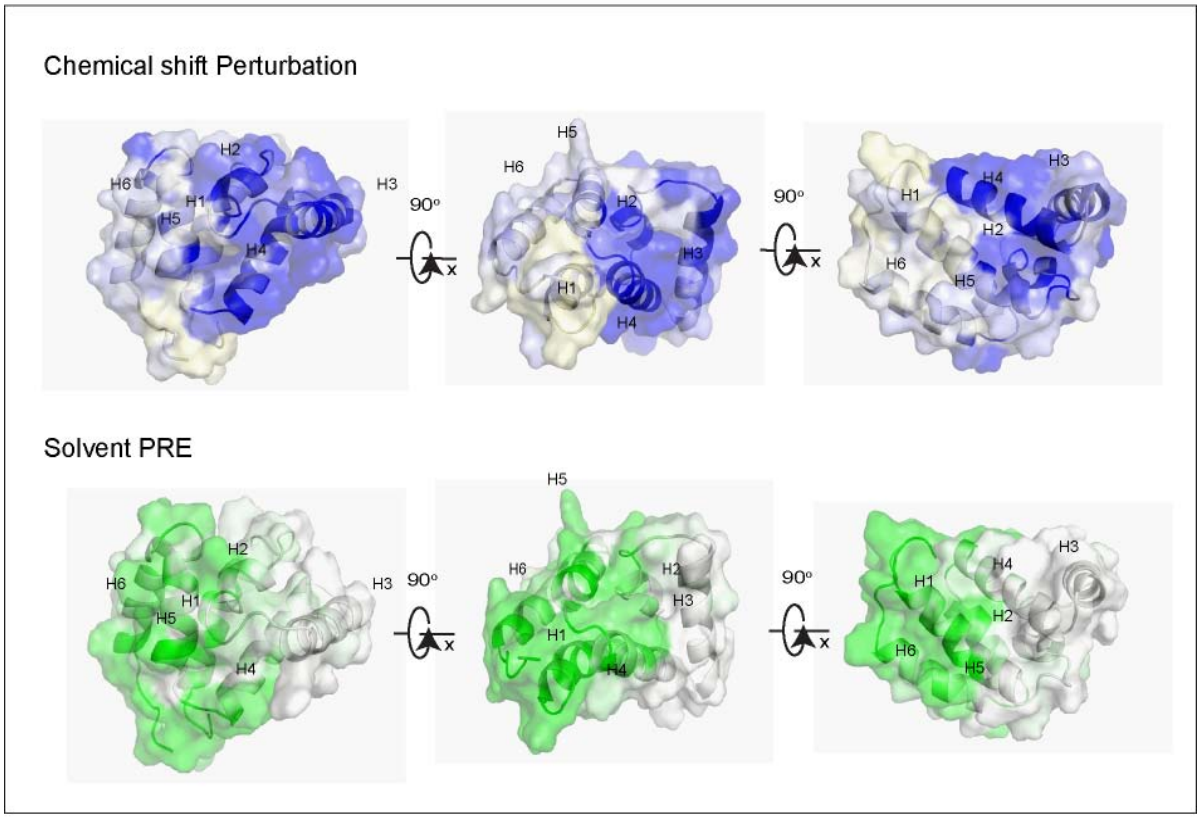

Figure 47. Surface representation of the residues having significant chemical shift perturbation and the residues having severe decrease of intensities in the presence of Gd-DPTA: The residues having significant chemical shift perturbation were involved into the assembly of the MAVS-$CARD_{1-100}$ filament indirectly or directly, while the residues bleached by Gd-DPTA were located on the surface of fibrillar assembly of MAVS-$CARD_{1-100}$.

in the absence and presence of 100 mM Gd-DTPA, respectively. The reason for recording two types of spectra was the overlapping peaks on one type of spectra would be resolved in the other type of spectra. For example, in the NCA spectra several alanine residues had severe overlap, but they were separated quite well on the ^{13}C-^{13}C correlation spectrum (**Figure 46**). The concentration of Gd-DTPA was around 5 times of that of the proteins in the solid state NMR rotor to make sure the proteins were surrounded by Gd-DTPA sufficiently. As shown in **Figure 46**, both carbon-carbon correlation spectra and NCA spectra recorded with 100 mM Gd-DTPA exhibited significant signal reduction on many residues. The residues with a severe drop of the peak volume (more than 50%) in the presence of Gd-DPTA were considered to be located on the surface of the MAVS-$CARD_{1-100}$ filament. In order to be more representative, these residues were plotted on the surface of the monomeric MAVS-$CARD_{1-100}$ (**Figure 47**). Notably, the surface plot of the residues bleached by Gd-DPTA was perfectly complementary to the surface plot of the residues exhibiting chemical shift perturbation, which indicated the sample were quite homogenous and provided a way for verifying the quaternary structure of MAVS-$CARD_{1-100}$ assembly.

3.4.3 Mutagenesis of the residues on the interaction surface

In order to confirm all interfaces, mutants on both sides of the three types of interaction surfaces were made and tested in reporter cell lines for monitoring the immune response. The experiments were carried out in collaboration with Dr. Renate König from the Paul-Ehrlich-Institut. Wild type MAVS and its mutants at different positions were transiently expressed in an ISRE-luc report cell line. The immune response caused by overexpression of MAVS and its mutants was monitored by a luciferase reporter assay. Western blot was also done to check the expression level of the transfected wt MAVS and its mutants with GAPDH as loading control. The results showed that all the mutants had significantly reduced activity to trigger the immune response, indicating all these residues were important for the activity of MAVS (**Figure 48**). Abrogating any of the interaction interfaces will cause the reduction of the immune response. In particular R64S R65S of type II interaction surface, R39S R41S of type III interaction surface and D23S E26S of type I interaction surface almost completely abolishing the stimulation of expression of interferon. These three double mutants D23S E26S, R39S R41S and R64S R65S were also found to disrupt the formation of the filament *in vitro* as judged by gel filtration chromatography (**Figure 22**). No intermediate complexes such as dimers

or trimers were observed, suggesting that filament formation might rely on the simultaneous presence of all the interfaces.

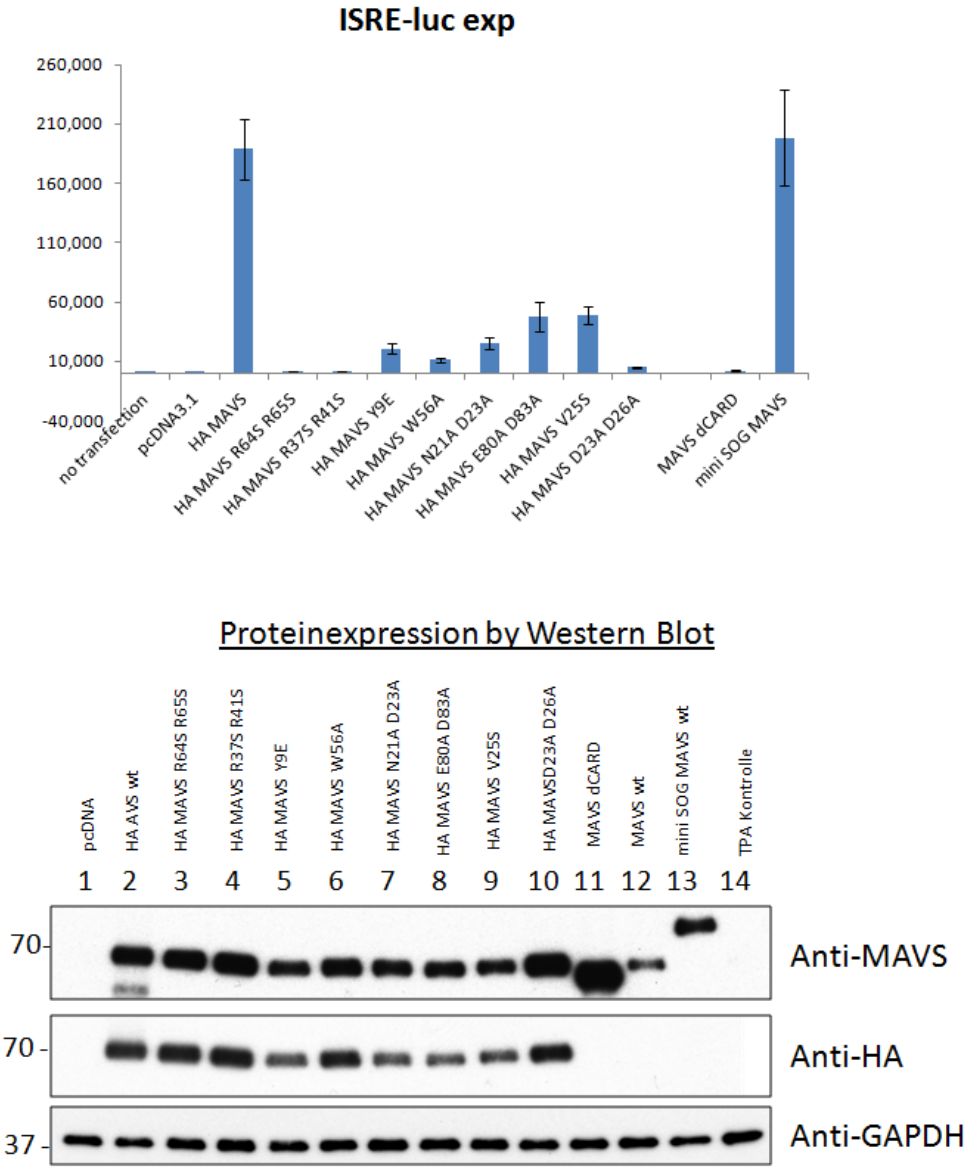

Figure 48. Bioactivity assay of MAVS and its mutants: MAVS protein with single or double mutations at all six interfaces were test in ISRE-luc report cell line for immune response. The expression level of the transfected MAVS or mutants were checked by western blot with GAPDH as loading control.

3.5 Quaternary structure of MAVS-CARD$_{1\text{-}100}$ filament

3.5.1 Symmetry information from electron microscopy images

Electron microscopic images of negatively stained specimens revealed that MAVS-CARD$_{1-100}$ formed high ordered structure in a filament state at pH 7.0 (**section 3.1.2**). To get further information, more negatively staining EM images were taken. Then the straight linear helical particles were selected using the EMAN2 software package (Ludtke et al., 1999; Rees et al., 2013). In total, 1971 non-redundant segments were selected. The power spectra of each individual segment was calculated and used to sort the images. In many cases, the images might contain fibers, but failed to diffract. The power spectrum provided a convenient way to discard these images. The similarity of the power spectra of most segments also indicated the sample was quite homogenous. The averaged power spectra in **Figure 49** showed a typical helical diffraction pattern. Normally, a helical object can be indexed by the Bessel order, n, and the layer line number, l. With the values of Bessel order and layer line number, the helical symmetry parameters: rise per subunit Δz and rotation per subunit $\Delta\Phi$ was estimated to be (6.3 Å, 140.8°) or (4.54 Å, 101.5°) based on the equations in **section 1.3.1.2**. The diameter of the fiber was estimated to be 82±2 Å from direct measurements on the EM images.

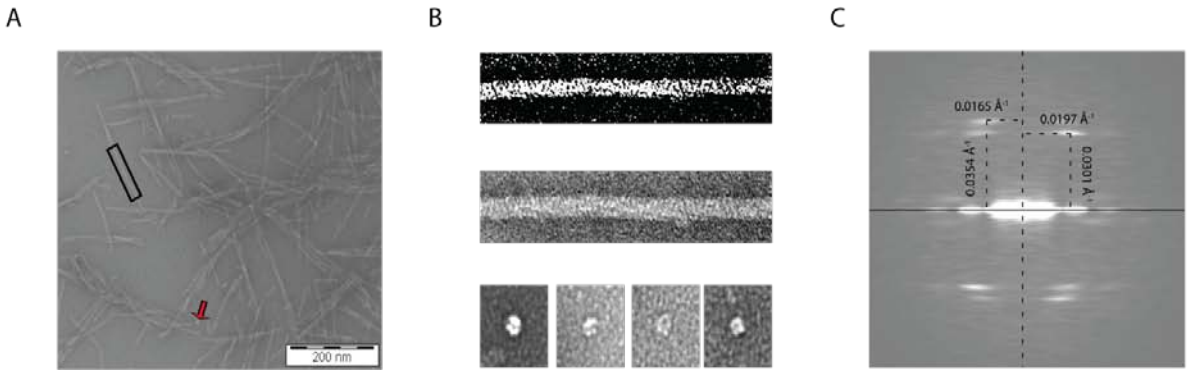

Figure 49. EM images of MAVS-CARD filament: A/B. Negative staining EM image of MAVS- CARD$_{1-100}$ filament. The boxed straight fiber is shown in B. The arrow indicates a small particle of MAVS- CARD$_{1-100}$ filament. Fortunately, several images of the cross section of the fiber were found after mild sonication and shown in B. **C.** The averaged electron diffraction pattern (power spectra) was calculated from 1971 non-redundant segments. Helix indexing was done based on the distance in the reciprocal space.

Later, higher resolution diffraction pattern were obtained in a collaboration with Dr. Sebastian Hiller *via* cyro EM. From the power spectra of cyro EM images, a clear layer line at 17.9 Å was visible that was not a meridional reflection (**Figure 50 A**). This layer line was considered to be a meridional reflection in the recently published paper from Xu. et. al (Xu et al., 2014). The Bessel order of 17.9 Å layer line was calculated to be 1 based on the **equation 1.12**. The distance of Bessel order 1 layer line corresponds to the pitch of the helix. The symmetry information Δz and $\Delta\Phi$ was then calculated to be (5.09 Å, 102.1°). The symmetry information could also be defined as 3.52 subunits per turn of 17.9 Å pitch. However, due to the intrinsic ambiguity of the EM power spectra, the precise symmetry information need to be refined by 3D reconstruction or calculated from NMR data. The class average of ~ 4000 segments shows the helical arrangement of MAVS-CARD$_{1\text{-}100}$ in the filament (**Figure 50 B**).

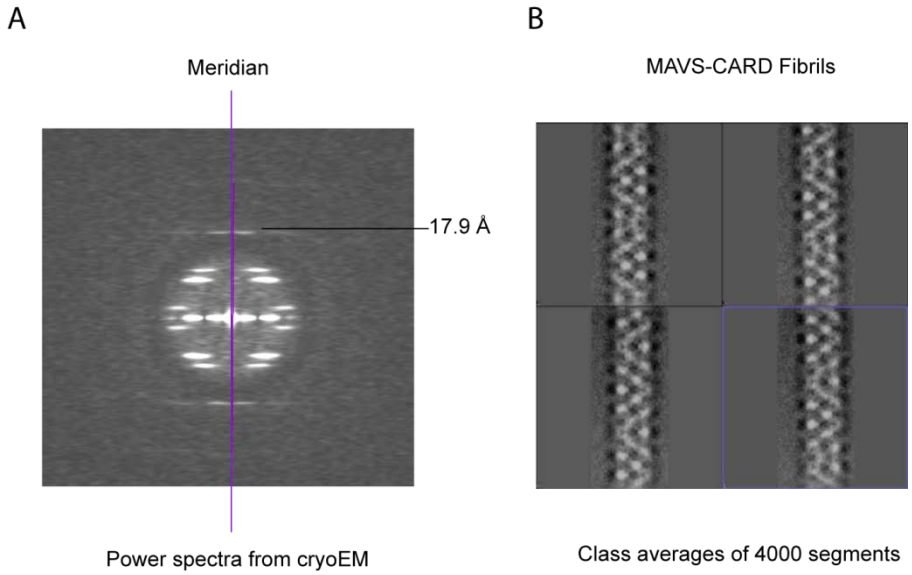

Figure 50. cryo EM of MAVS-CARD fibrils:. **A.** The averaged power spectra from cyro EM images. **B.** Class average of 4000 segments shows the helical assembly of MAVS-CARD$_{1\text{-}100}$.

3.5.2 Quaternary structure of MAVS-CARD$_{1\text{-}100}$ filament calculated by ARIA

The quaternary structure calculation is a challenging task due to various, often unknown symmetries of the protein complexes and lacking of matured software tools to handle this symmetry problem. In this project, the quaternary structure calculation was carried out in collaboration with Dr. Benjamin

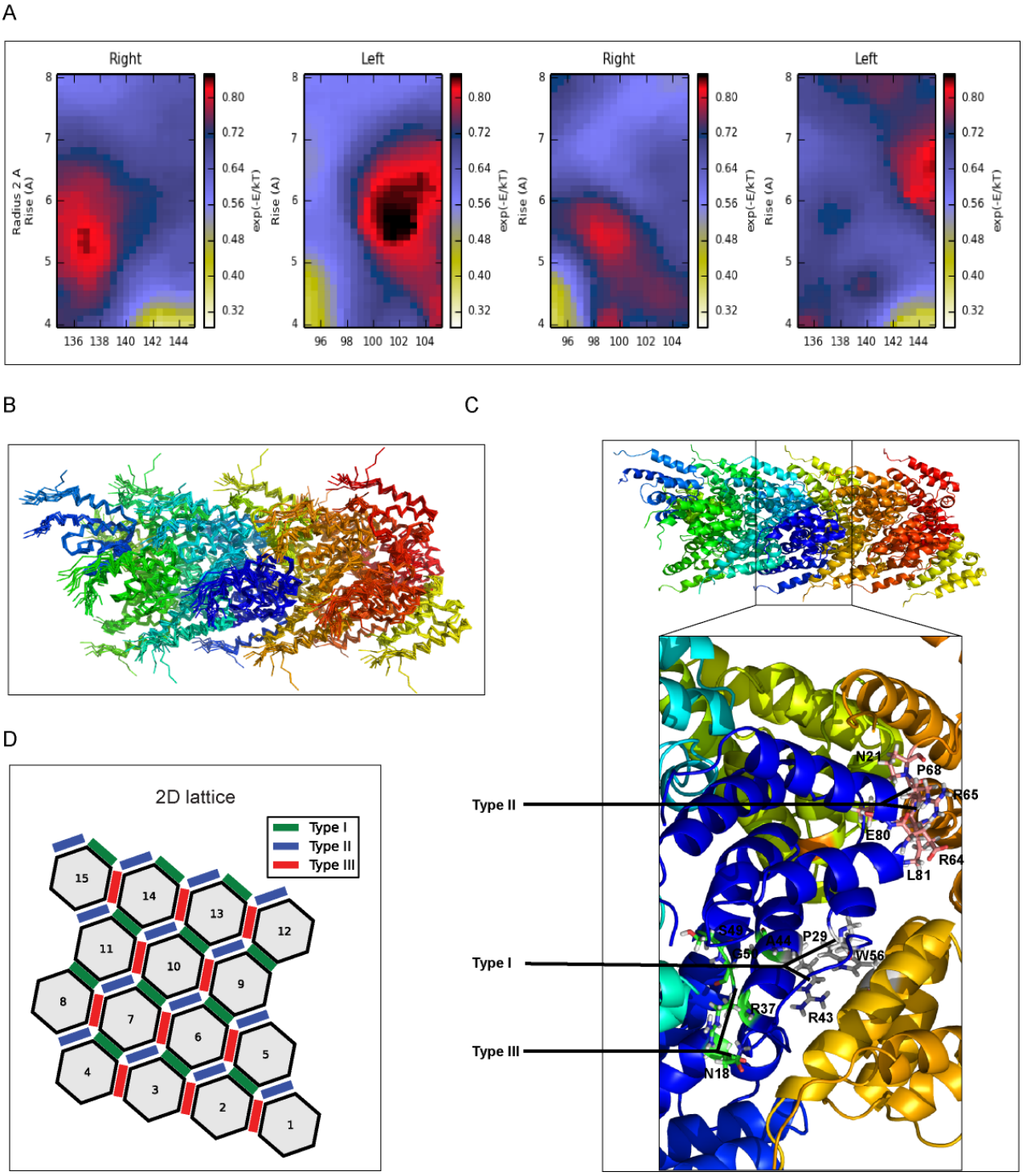

Figure 51: Quaternary structure of MAVS-CARD$_{1\text{-}100}$ filament calculated by ARIA using NMR constraints. A. Grid search based on the unambiguously assigned inter-molecular constraints reveals a helical symmetry of MAVS-CARD1-100 with azimuthal rotation angle 101.5°, axial translation 5.2 Å and the handness is determined to be left-handed . **B.** 10 lowest-energy structures calculated by ARIA with all the intra- and inter-molecular constraints. **C.** Cartoon diagram of the structure of MAVS-CARD$_{1\text{-}100}$ filament. The heptamer consisting of one MAVS-CARD$_{1\text{-}100}$

subunit surrounded by six identical nearest neighbors is shown with the three type interaction interfaces. **D.** 2D lattice represented all three types interaction in the context of MAVS-CARD$_{1-100}$ filament.

Bardiaux. First, he performed a grid search using the unambiguous intermolecular constraints to get the symmetry information. It was done by scanning azimuthal rotation angle ($\Delta\Phi$) from 94° to 106° degrees and from 134° to 146°, axial translation (Δz) from 4 Å to 8 Å, with both right and left handness to look for the minimum energy status with the calculated protomer structure and intermolecular constraints. In **Figure 51 A**, the energy distritbutions for the ten lowest energy structures of each grid search are shown. A clear energy minimum is apparent for the symmetry with azimuthal rotation angle of 101.5°, an axial translation of 5.2 Å and left handness, which is consistent with the symmetry information derived from EM images. Then the helical symmetry was incorporated into the software ARIA. With all intra- and inter- molecular constraints, the quaternary structure of MAVS-CARD$_{1-100}$ filament was calculated. 10 lowest-energy structures are shown in **Figure 51 B**. Each MAVS-CARD$_{1-100}$ subunit interacted with the six spatially closest neighbors through the three type of interaction interfaces, and all protomers were in the identical environment regardless of their position in the filament (**Figure 51 C**). 2D lattice represented all three types interaction in the context of MAVS-CARD1-100 filament was shown in **Figure 51 D**.

4 Discussion and Perspectives

4.1 MAVS location on the outer membrane of the mitochondria

Mitochondria are dynamic, double membrane cellular organelles that are involved in a wide range of critical cellular processes, including cell apoptosis, ATP generation and calcium homeostasis. However, more and more evidence suggests that mitochondria also participate in various innate immune signaling pathways. MAVS as a mitochondria-anchored adaptor protein was found to mediate RLR signaling in a prion-like way through autocatalytical aggregation of its CARDs (Hou et al., 2011). However, it is unclear how MAVS could keep monomeric in its inactive state. Buffer condition screening of MAVS-CARD filament reveals its self-assembly can be disrupted at lower pH (**Figure 16**). Considering the proton gradient across the mitochondria inner membrane and the free permission of protons across the outer membrane, the pH directly at the outer membrane is lower than that in the cytosol and the mitochondrial matrix (Porcelli et al., 2005). Even though the precise value of the pH around the outer membrane of mitochondria is still unknown, the lower pH could help to stabilize MAVS in its inactive state. For CARD mediated assembly of MAVS, it is essential to have most of, if not all, the interfaces available, mutagenesis on each of the three types of interaction interfaces could disrupt the formation of the filament assembly. Comparison of the gel filtration results of MAVS $_{1\text{-}100}$ and other truncated versions of MAVS such as MAVS $_{1\text{-}175}$, MAVS $_{1\text{-}500}$ showed the latter two truncated versions of MAVS formed smaller aggregates, indicating the CARD domain of MAVS might interact with other parts of full length MAVS to keep itself in a 'locked' state when it is inactive (**Figure 14**). Since the six interaction surfaces cover a large area of the surface of MAVS-CARD$_{1\text{-}100}$, the region 100-500 of MAVS might shield one of the interfaces, if there is interaction, blocking the self assembly of full length MAVS in its inactive state.

4.2 RIG-I CARD might serve as a nucleator for MAVS self-assembly

In this project, RIG-I CARD $_{1\text{-}172}$ was shown to interact with MAVS-CARD$_{1\text{-}100}$ via both pull down assay and titration assay using solution NMR without adding of poly-ubiquitin chains. Previous studies showed K63 ubiquitin chains were needed in the catalytic activation of MAVS assembly by RIG-I (Zeng et al., 2010). However, this was not consistent. Peisley and coworkers also reported

RIG-I formed a filament along dsRNA which could promote 2CARD oligomerization and stimulate MAVS activation without K63 ubiquitin chains (Peisley et al., 2013). Later, another paper revealed that the ubiquitin chains bridged adjacent 2CARD of RIG-I and stabilized the tetramer of 2CARD of RIG-I. A simple release of 2CARD was insufficient for the signal activation of MAVS CARD. Adding K63 ubiquitin chains could promote 2CARD tetramerization itself (Peisley et al., 2014). For the RIG-I $2CARD_{1-172}$ studied in this project, gel filtration of it revealed it contained both high molecular weight oligomer and monomer (**Figure 14**). The oligomer could be the effective component to interact with MAVS-CARD without K63 ubiquitin chains. The next question is how the oligomeric, not monomeric, 2CARD could activate MAVS CARD filament formation. In the case of other death domain complexes like PIDDosome and Myddosome, the upstream helical oligomeric signalling complex served as a scaffold for downstream molecules containing death domains. RIG-I 2CARD oligomer was speculated to activate MAVS CARD in a similar way through extension of the helical complex. The titration assay using solution NMR revealed a large interaction surface of RIG-I 2CARD on MAVS CARD. In **Figure 52**, the residues of MAVS CARD involved in interaction with RIG-I are colored in blue. If this molecule were overlaid with another MAVS CARD molecule in the context of three interaction surfaces, the blue parts matched the six interfaces quite well (**Figure 52**). The comparison between the structure of the second CARD domain of RIG-I (RIG

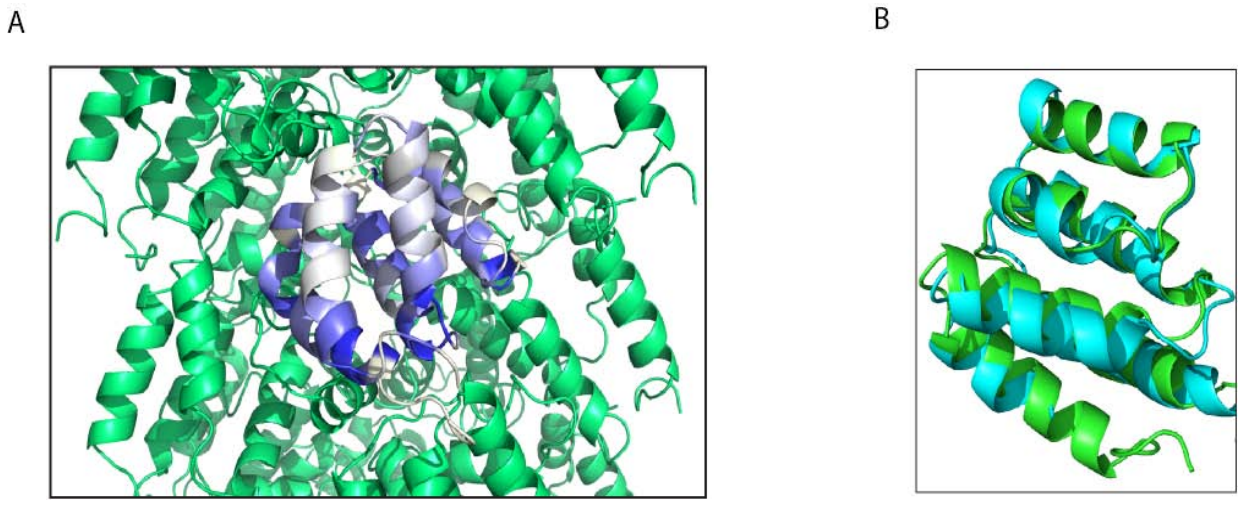

Figure 52. RIG-I 2CARDs serve as a nucleator for MAVS self-association:. **A.** Labeling of RIG-I 2CARD interaction sites on MAVS CARD in the context of three types interactions. **B.** Overlay of MAVS CARD with second CARD of RIG-I. They have quite similar structure with r.m.s.d. 1.042 Å. MAVS CARD is in green and the second CARD of RIG-I is in cyan.

-I CARD2) with MAVS CARD revealed they were quite similar although they had only 15% sequence identity (**Figure 52**). Furthermore, the recently published crystal structure of RIG-I 2CARD exhibited its tetramer was mediated through an analogous six interfaces for both 2 CARDs (type II interaction) and for intermolecular interactions (type I and type III) in the tetramer (Peisley et al., 2014). All these data are consistent with the notion that the RIG-I 2CARD serves as a nucleator for MAVS assembly.

4.3 Characterization of MAVS-CARD filament

The MAVS CARD filament has several features similar to prions, such as protease and detergent resistance and self-perpetuation (Hou et al., 2011). However, MAVS CARD filament cannot be stained by Congo red and is lacking glutamine and aspargine-rich regions, both of which are typical characters of amyloid fibers. The CD spectrum shows that MAVS CARD secondary structure is mainly alpha-helical. The structure of a MAVS CARD protomer within its filament reveals it is almost the same as the structure of monomeric MAVS CARD. MAVS CARDs form the filament assembly based on the collective interactions at the six interfaces of each subunit (**Figure 45**). The difference of orientation of helix3 is probably due to the participation of helix 3, as well as helix 2-helix 3 and helix 3- helix 4 loops in all three types of interaction in the filament (**Figure 45**). Among known DD assemblies, ASC^{PYD} with its helical assembly possesses a similar orientational change of helix2- helix3 loops and helix 3 compared to the structure of ASC^{PYD} alone, suggesting such a conformational change might facilitate the formation of the helical assembly (Lu et al., 2014). Notably, one face of the cross section of a MAVS-$CARD_{1-100}$ filament is highly positively charged and the opposite one is mostly negatively charged (**Figure 53**), which indicates that the charge complementarity plays an importance role in the filament assembly. This is consistent with the experiment that the filament could be disrupted into monomer in pH 3.0 buffer. In contrast, only one cross section of RIG-I 2CARD tetramer shows a clearly negatively charged surface, which is on the side of the second CARD domain. The other side (CARD1) contains almost equivalently positively and negatively charged residues. As discussed in the last section, CARD2 shared a lot of similarity with MAVS CARD. Likewise, a recent publication showed that mutagenesis on CARD2, but not on CARD1 of RIG-I significantly affected the signaling activity of RIG-I, suggesting CARD2 was involved in MAVS recruitment (Peisley et al., 2014). Based on these observations, a model of nucleation of MAVS CARD by RIG-I 2CARD is proposed in **Figure 53**. The CARD2s of RIG-I recruit MAVS-CARD mainly through electrostatic interactions for helical extension (**Figure 53**).

Due to a similar negatively charged cross section on one side of MAVS CARD filament, the prion like behavior of MAVS CARD might share the same mechanism as RIG-I 2CARD activation of MAVS. The filament formation of MAVS CARDs relies on the collective interactions at six interfaces of each molecule. The tight packing of the filament and relatively small surface to volume

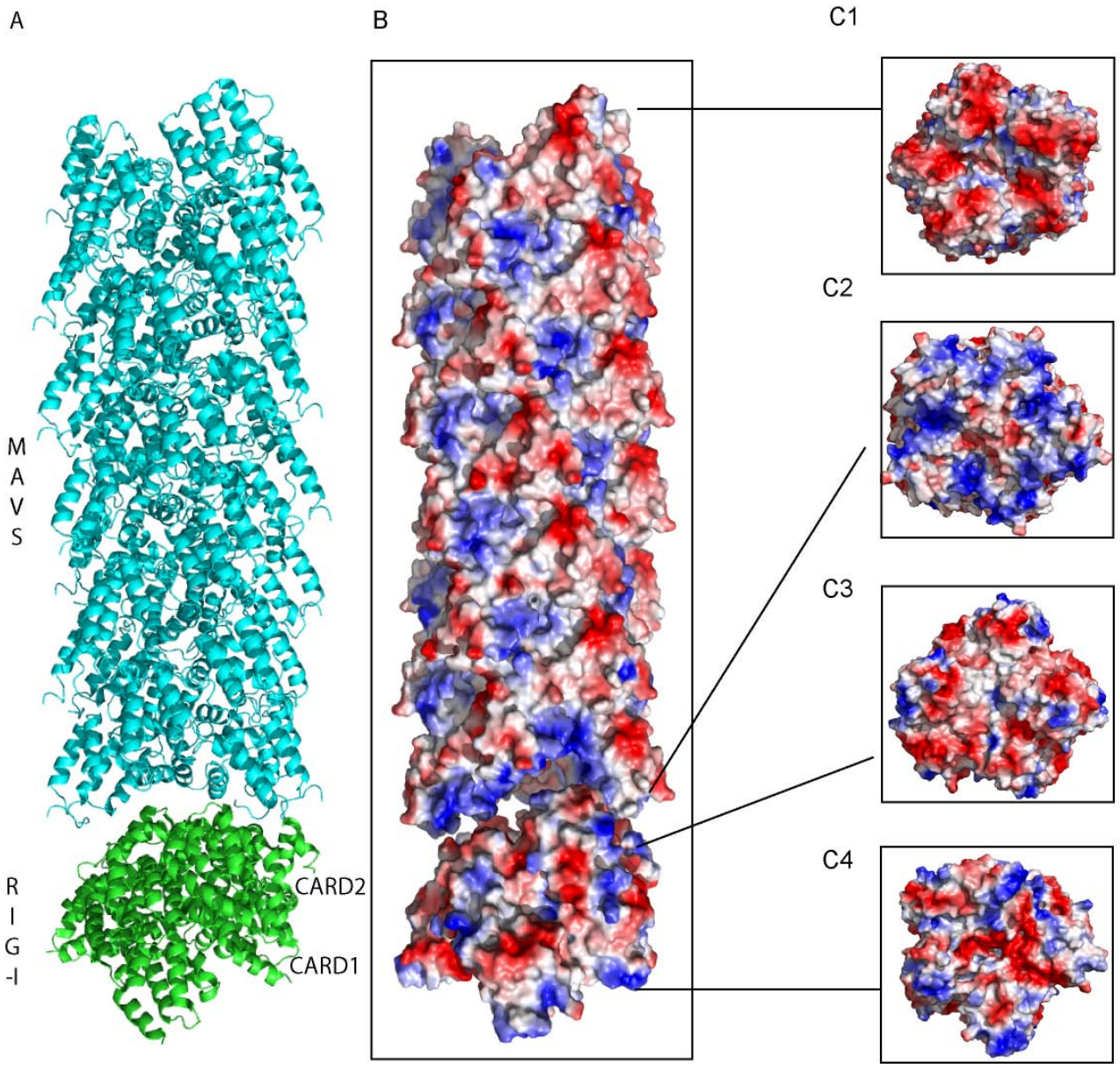

Figure 53. Model of the nucleation of MAVS CARD by RIG-I 2CARD: A. Cartoon representation of the model for RIG-I 2CARD tetramer nucleating MAVS CARD. The tetramer of RIG-I 2CARD is in green. MAVS CARD filament is in cyan. **B.** Electrostatic potential of MAVS CARD filament and RIG-I 2CARD tetramer. **C.** The electrostatic potential at the cross sections of MAVS CARD filament and RIG-I 2CARD tetramer. MAVS filament has negatively charged cross

section at one side (C1) and positively charged surface on the other side (C2). RIG-I 2CARD tetramer has only one negatively charged cross section on the CARD2 side (C3).

ratio probably make the filaments detergent and protease-resistant. Thus the longer MAVS filaments would entail a more stable amplification of the innate immune signaling. However, it is still needed to be studied how such higher-order complexes degrade or disassemble in cells for terminating the innate immune signaling.

4.4 The oligomerization of MAVS CARD is important for efficient immune response

The atomic model of MAVS-$CARD_{1-100}$ filament from the combination of NMR, EM and X-ray fiber diffraction reveals that MAVS CARD assembled into a helical filament without any major secondary structure conformational change. It is surprising to have filament growth of a membrane-anchored protein. However, membrane anchoring is important for the activity of full length MAVS. Secondary structure prediction suggests that the ~ 400 amino acids segment between CARD domain and TM domain is random coil, providing large flexibility for CARD domain to form a fibrillar higher order structure (**Figure 4 C**). Notably, this random coil region is supposed to interact with multiple down-streaming signaling molecules (Barbalat et al., 2011; Castanier et al., 2010) upon on aggregation of MAVS. Thus the helical arrangement of CARD domains and the anchoring of the TM domain on the mitochondria seem to provide a good solution to enable the random coil region having less extended conformation, facilitating efficient binding of its interaction partners. On the other hand, the higher-order oligomerization might be required for avidity-driven interactions or creating new binding sites. The assembly of MAVS into a helical filament could reduce the 'off-rate' of possible low affinity interactions between MAVS and its down-streaming signaling components. The MAVS polymer is like a semisolid interaction matrix that would be expected to have different diffusion kinetics with traditional bio-molecules. Such semisolid matrix-mediated interactions have been described for clathrin (Schmid et al., 2006) and E4orf3 (Ou et al., 2012), where multiple weak binding sites form a stronger chelating surface after polymerization. It is the case that the assembly of MAVS is essential for interaction with TRAF, RIP1 and FADD (Takeuchi and Akira, 2009; Tang and Wang, 2009). The long MAVS filament would enable a significant amplification of the innate immune signal. Because both MAVS downstream signalling components TRAF2 and TRAF3 form trmers, the MAVS filament could enhance their interaction with MAVS via increased avidity (Napetschnig and Wu,

2013; Saha et al., 2006; Zheng et al., 2010). This reveals a paradigm whereby the individual proteins assemble to form higher-order polymers serving as an interaction matrix to bind with interaction partners and activate the downstream signaling. Moreover, multiple contacts in a flexible scaffold have also been demonstrated to be more specific and efficient than the stronger individual interactions (Martos et al., 2008), which is important for immune response to be fast and precisely controlled in host cells.

4.5 High-ordered assemblies represent a new paradigm of signal transduction.

More and more studies in innate immunity and apoptosis pathways revealed multiple signaling proteins assembling into higher order aggregates for transmission or amplification of receptor activation information into cellular responses. For example, the Apaf1 apoptosome and the PIDDosome are involved in cell apoptosis. DISC complex and ASC^{PYD} filament played an important role in innate immune response. The oligomeric assemblies of RIG-I 2CARD and MAVS CARD reinforce the new paradigm of signal transduction via higher-order signaling complexes (Park et al., 2007b; Peisley et al., 2014; Wang et al., 2010) for signal amplification. All of these signaling complexes are assembled through members of the DD superfamily. The study of the structural basis for MAVS activation in this project showed the six interfaces in the MAVS CARD filament and revealed the mechanism of MAVS CARD prion like behavior. The structural organization of MAVS filament is different from β-amyloid prions. A model of RIG-I 2CARD complex triggering MAVS CARD activation is proposed based on the structure of MAVS CARD filament and the interaction study of RIG-I 2CARD with MAVS CARD mutants. The polymerization of MAVS CARD without secondary structure conformational change facilitates a robust and sensitive immune response against infections. Further *in vivo* studies on MAVS aggregation and distribution on mitochondria, its correlation with other mitochondrial proteins and mitochondrial dynamics will shed more light on the process of MAVS mediate immune response.

5 Reference

Ahmed, M.A., Bamm, V.V., Harauz, G., and Ladizhansky, V. (2007). The BG21 isoform of Golli myelin basic protein is intrinsically disordered with a highly flexible amino-terminal domain. Biochemistry *46*, 9700-9712.

Akira, S., Uematsu, S., and Takeuchi, O. (2006). Pathogen recognition and innate immunity. Cell *124*, 783-801.

Al-Khayat, H.A., Morris, E.P., Kensler, R.W., and Squire, J.M. (2008). Myosin filament 3D structure in mammalian cardiac muscle. Journal of structural biology *163*, 117-126.

Arimoto, K., Takahashi, H., Hishiki, T., Konishi, H., Fujita, T., and Shimotohno, K. (2007). Negative regulation of the RIG-I signaling by the ubiquitin ligase RNF125. Proc Natl Acad Sci U S A *104*, 7500-7505.

Arnoult, D., Carneiro, L., Tattoli, I., and Girardin, S.E. (2009). The role of mitochondria in cellular defense against microbial infection. Seminars in immunology *21*, 223-232.

Baldus, M., and Meier, B.H. (1996). Total correlation spectroscopy in the solid state. The use of scalar couplings to determine the through-bond connectivity. J Mag Res A *121*, 65-69.

Barbalat, R., Ewald, S.E., Mouchess, M.L., and Barton, G.M. (2011). Nucleic acid recognition by the innate immune system. Annu Rev Immunol *29*, 185-214.

Bardiaux, B., Bernard, A., Rieping, W., Habeck, M., Malliavin, T.E., and Nilges, M. (2008). Graphical analysis of NMR structural quality and interactive contact map of NOE assignments in ARIA. BMC Struct Biol *8*, 30.

Baril, M., Racine, M.E., Penin, F., and Lamarre, D. (2009). MAVS dimer is a crucial signaling component of innate immunity and the target of hepatitis C virus NS3/4A protease. J Virol *83*, 1299-1311.

Bassolino-Klimas, D., Tejero, R., Krystek, S.R., Metzler, W.J., Montelione, G.T., and Bruccoleri, R.E. (1996). Simulated annealing with restrained molecular dynamics using a flexible restraint potential: theory and evaluation with simulated NMR constraints. Protein science : a publication of the Protein Society *5*, 593-603.

Bernini, A., Spiga, O., Venditti, V., Prischi, F., Bracci, L., Tong, A.P., Wong, W.T., and Niccolai, N. (2006). NMR studies of lysozyme surface accessibility by using different paramagnetic relaxation probes. Journal of the American Chemical Society *128*, 9290-9291.

Biasini, E., Turnbaugh, J.A., Unterberger, U., and Harris, D.A. (2012). Prion protein at the crossroads of physiology and disease. Trends in neurosciences *35*, 92-103.

Blanchet, C.E., and Svergun, D.I. (2013). Small-angle X-ray scattering on biological macromolecules and nanocomposites in solution. Annual review of physical chemistry *64*, 37-54.

Botos, I., Liu, L., Wang, Y., Segal, D.M., and Davies, D.R. (2009). The toll-like receptor 3:dsRNA signaling complex. Biochim Biophys Acta *1789*, 667-674.

Boyd, J., Hommel, U., and Campbell, I.D. (1990). Influence of cross-correlation between dipolar and anisotropic chemical shift relaxation mechanisms upon longitudinal relaxation rates of 15N in macromolecules. Chem Phys Lett *175*, 477-482.

Brateman, L. (1986). Chemical shift imaging: a review. AJR American journal of roentgenology *146*, 971-980.

Broquet, A.H., Hirata, Y., McAllister, C.S., and Kagnoff, M.F. (2011). RIG-I/MDA5/MAVS are required to signal a protective IFN response in rotavirus-infected intestinal epithelium. J Immunol *186*, 1618-1626.

Brunger, A.T. (2007). Version 1.2 of the Crystallography and NMR system. Nature protocols *2*, 2728-2733.

Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., *et al.* (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr D Biol Crystallogr *54*, 905-921.

Cardenas, W.B., Loo, Y.M., Gale, M., Jr., Hartman, A.L., Kimberlin, C.R., Martinez-Sobrido, L., Saphire, E.O., and Basler, C.F. (2006). Ebola virus VP35 protein binds double-stranded RNA and inhibits alpha/beta interferon production induced by RIG-I signaling. J Virol *80*, 5168-5178.

Castanier, C., Garcin, D., Vazquez, A., and Arnoult, D. (2010). Mitochondrial dynamics regulate the RIG-I-like receptor antiviral pathway. EMBO reports *11*, 133-138.

Castanier, C., Zemirli, N., Portier, A., Garcin, D., Bidere, N., Vazquez, A., and Arnoult, D. (2012). MAVS ubiquitination by the E3 ligase TRIM25 and degradation by the proteasome is involved in type I interferon production after activation of the antiviral RIG-I-like receptors. BMC biology *10*, 44.

Castellani, F., van Rossum, B., Diehl, A., Schubert, M., Rehbein, K., and Oschkinat, H. (2002). Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. Nature *420*, 98-102.

Cavanagh, J., Fairbrother, W.J., Palmer, A.G., Skelton, N.J., and Rance, M. (2007). Protein NMR Spectroscopy, Principles and Practice, 2 edn (Amsterdam: Academic Press).

Clare, D.K., and Orlova, E.V. (2010). 4.6A Cryo-EM reconstruction of tobacco mosaic virus from images recorded at 300 keV on a 4k x 4k CCD camera. Journal of structural biology *171*, 303-308.

Cole, C., Barber, J.D., and Barton, G.J. (2008). The Jpred 3 secondary structure prediction server. Nucleic acids research *36*, W197-201.

Culf, A.S., Cuperlovic-Culf, M., and Ouellette, R.J. (2009). Chemical exchange in novel spirobicyclic zwitterionic Janovsky complexes using dynamic 1H NMR spectroscopy. Magnetic resonance in chemistry : MRC *47*, 158-164.

Cuperlovic, M., Meresi, G.H., Palke, W.E., and Gerig, J.T. (2000). Spin relaxation and chemical exchange in NMR simulations. Journal of magnetic resonance *142*, 11-23.

Dalgarno, D.C., Levine, B.A., and Williams, R.J. (1983). Structural information from NMR secondary chemical shifts of peptide alpha C-H protons in proteins. Bioscience reports *3*, 443-452.

De Rosier, D.J., and Klug, A. (1968). Reconstruction of three dimensional structures from electron micrographs. Nature *217*, 130-134.

Diaz, R., Rice, W.J., and Stokes, D.L. (2010). Fourier-Bessel reconstruction of helical assemblies. Methods Enzymol *482*, 131-165.

Dixit, E., Boulant, S., Zhang, Y., Lee, A.S., Odendall, C., Shum, B., Hacohen, N., Chen, Z.J., Whelan, S.P., Fransen, M., *et al.* (2010). Peroxisomes are signaling platforms for antiviral innate immunity. Cell *141*, 668-681.

Doreleijers, J.F., Sousa da Silva, A.W., Krieger, E., Nabuurs, S.B., Spronk, C.A., Stevens, T.J., Vranken, W.F., Vriend, G., and Vuister, G.W. (2012). CING: an integrated residue-based structure validation program suite. J Biomol NMR *54*, 267-283.

Dosset, P., Hus, J.C., Blackledge, M., and Marion, D. (2000). Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data. J Biomol NMR *16*, 23-28.

Dusold, S., and Sebald, A. (2000). Dipolar Recoupling under Magic-Angle-Spinning Conditions. In Annal Rep NMR Spec, G.A. Webb, ed. (Elsevier), pp. 185-264.

Egelman, E.H. (2007). The iterative helical real space reconstruction method: surmounting the problems posed by real polymers. Journal of structural biology *157*, 83-94.

Egelman, E.H. (2010). Reconstruction of helical filaments and tubes. Methods Enzymol *482*, 167-183.

Ernst, R.R., Bodenhausen, G., and Wokaun, A. (1987). Principles of Nuclear Magnetic Resonance in One and Two Dimensions. (Oxford: Clarendon Press).

Evans, J.N.S. (1995). Biomolecular NMR Spectroscopy (New York: Oxford).

Farrow, N.A., Zhang, O.W., Formankay, J.D., and Kay, L.E. (1995). Comparison of the backbone dynamics of a folded and an unfolded SH3 domain existing in equilibrium in aqueous buffer. Biochemistry *34*, 868-878.

Faul, E.J., Wanjalla, C.N., Suthar, M.S., Gale, M., Wirblich, C., and Schnell, M.J. (2010). Rabies virus infection induces type I interferon production in an IPS-1 dependent manner while dendritic cell activation relies on IFNAR signaling. PLoS Pathog *6*, e1001016.

Ferrao, R., and Wu, H. (2012). Helical assembly in the death domain (DD) superfamily. Current opinion in structural biology *22*, 241-247.

Franks, W.T., Wylie, B.J., Schmidt, H.L., Nieuwkoop, A.J., Mayrhofer, R.M., Shah, G.J., Graesser, D.T., and Rienstra, C.M. (2008). Dipole tensor-based atomic-resolution structure determination of a nanocrystalline protein by solid-state NMR. Proc Natl Acad Sci U S A *105*, 4621-4626.

Gack, M.U., Albrecht, R.A., Urano, T., Inn, K.S., Huang, I.C., Carnero, E., Farzan, M., Inoue, S., Jung, J.U., and Garcia-Sastre, A. (2009). Influenza A virus NS1 targets the ubiquitin ligase TRIM25 to evade recognition by the host viral RNA sensor RIG-I. Cell host & microbe *5*, 439-449.

Gack, M.U., Shin, Y.C., Joo, C.H., Urano, T., Liang, C., Sun, L., Takeuchi, O., Akira, S., Chen, Z., Inoue, S., *et al.* (2007). TRIM25 RING-finger E3 ubiquitin ligase is essential for RIG-I-mediated antiviral activity. Nature *446*, 916-920.

Gardner, K.H., and Kay, L.E. (1998). The use of ^{2}H, ^{13}C, ^{15}N multidimensional NMR to study the structure and dynamics of proteins. Annual review of biophysics and biomolecular structure *27*, 357-406.

Griffin, R.G. (1998). Dipolar recoupling in MAS spectra of biological solids. Nature structural biology *5 Suppl*, 508-512.

Guerry, P., and Herrmann, T. (2012). Comprehensive automation for NMR structure determination of proteins. Methods Mol Biol *831*, 429-451.

Guntert, P. (2004). Automated NMR structure calculation with CYANA. Methods Mol Biol *278*, 353-378.

Guntert, P., Mumenthaler, C., and Wuthrich, K. (1997). Torsion angle dynamics for NMR structure calculation with the new program DYANA. J Mol Biol *273*, 283-298.

Hayakawa, S., Shiratori, S., Yamato, H., Kameyama, T., Kitatsuji, C., Kashigi, F., Goto, S., Kameoka, S., Fujikura, D., Yamada, T., *et al.* (2011). ZAPS is a potent stimulator of signaling mediated by the RNA helicase RIG-I during antiviral responses. Nat Immunol *12*, 37-44.

Heil, F., Hemmi, H., Hochrein, H., Ampenberger, F., Kirschning, C., Akira, S., Lipford, G., Wagner, H., and Bauer, S. (2004). Species-specific recognition of single-stranded RNA via toll-like receptor 7 and 8. Science *303*, 1526-1529.

Herrmann, T., Guntert, P., and Wuthrich, K. (2002). Protein NMR structure determination with automated NOE-identification in the NOESY spectra using the new software ATNOS. J Biomol NMR *24*, 171-189.

Hohwy, M., Rienstra, C.M., and Griffin, R.G. (2002). Band-selective homonuclear dipolar recoupling in rotating solids. J Chem Phys *117*, 4973-4987.

Hohwy, M., Rienstra, C.M., Jaroniec, C.P., and Griffin, R.G. (1999). Fivefold symmetric homonuclear dipolar recoupling in rotating solids: Application to double quantum spectroscopy. J Chem Phys *110*, 7983-7992.

Holm, L., and Rosenstrom, P. (2010). Dali server: conservation mapping in 3D. Nucleic acids research *38*, W545-549.

Hong, M. (1999). Determination of multiple ***φ***-torsion angles in proteins by selective and extensive (13)C labeling and two-dimensional solid-state NMR. Journal of magnetic resonance *139*, 389-401.

Hong, M., and Jakes, K. (1999). Selective and extensive 13C labeling of a membrane protein for solid-state NMR investigations. J Biomol NMR *14*, 71-74.

Hou, F., Sun, L., Zheng, H., Skaug, B., Jiang, Q.X., and Chen, Z.J. (2011). MAVS forms functional prion-like aggregates to activate and propagate antiviral innate immune response. Cell *146*, 448-461.

Ishikawa, H., and Barber, G.N. (2008). STING is an endoplasmic reticulum adaptor that facilitates innate immune signalling. Nature *455*, 674-678.

Iwasaki, A., and Medzhitov, R. (2010). Regulation of adaptive immunity by the innate immune system. Science *327*, 291-295.

Janssens, S., Tinel, A., Lippens, S., and Tschopp, J. (2005). PIDD mediates NF-kappaB activation in response to DNA damage. Cell *123*, 1079-1092.

Jardetsky, O., and Roberts, G.C.K. (1981). NMR in Molecular Biology (New York: Academic Press).

Jaroniec, C.P., MacPhee, C.E., Bajaj, V.S., McMahon, M.T., Dobson, C.M., and Griffin, R.G. (2004). High-resolution molecular structure of a peptide in an amyloid fibril determined by magic angle spinning NMR spectroscopy. Proc Nat Acad Sci USA *101*, 711-716.

Jia, Y., Song, T., Wei, C., Ni, C., Zheng, Z., Xu, Q., Ma, H., Li, L., Zhang, Y., He, X., *et al.* (2009). Negative regulation of MAVS-mediated innate immune response by PSMA7. J Immunol *183*, 4241-4248.

Jounai, N., Takeshita, F., Kobiyama, K., Sawano, A., Miyawaki, A., Xin, K.Q., Ishii, K.J., Kawai, T., Akira, S., Suzuki, K., *et al.* (2007). The Atg5 Atg12 conjugate associates with innate antiviral immune responses. Proc Natl Acad Sci U S A *104*, 14050-14055.

Kanelis, V., Forman-Kay, J.D., and Kay, L.E. (2001). Multidimensional NMR methods for protein structure determination. IUBMB life *52*, 291-302.

Karplus, M. (1959). Contact Electron-Spin Coupling of Nuclear Magnetic Moments. J Chem Phys *30*, 11-15.

Karplus, M. (1963). Vicinal Proton Coupling in Nuclear Magnetic Resonance. J Am Chem Soc *85*, 2870-2871.

Kato, H., Takeuchi, O., Sato, S., Yoneyama, M., Yamamoto, M., Matsui, K., Uematsu, S., Jung, A., Kawai, T., Ishii, K.J., *et al.* (2006). Differential roles of MDA5 and RIG-I helicases in the recognition of RNA viruses. Nature *441*, 101-105.

Katze, M.G., He, Y., and Gale, M., Jr. (2002). Viruses and interferon: a fight for supremacy. Nat Rev Immunol *2*, 675-687.

Kawai, T., and Akira, S. (2009). The roles of TLRs, RLRs and NLRs in pathogen recognition. Int Immunol *21*, 317-337.

Kawai, T., Takahashi, K., Sato, S., Coban, C., Kumar, H., Kato, H., Ishii, K.J., Takeuchi, O., and Akira, S. (2005). IPS-1, an adaptor triggering RIG-I- and Mda5-mediated type I interferon induction. Nat Immunol *6*, 981-988.

Kay, L.E. (1998). Protein dynamics from NMR. Nature structural biology *5 Suppl*, 513-517.

Kay, L.E. (2005). NMR studies of protein structure and dynamics. J Mag Res *173*, 193-207.

Kay, L.E., Torchia, D.A., and Bax, A. (1989). Backbone dynamics of proteins as studied by N-15 inverse detected heteronuclear NMR-spectroscopy - Application to staphylococcal nuclease. Biochemistry *28*, 8972-8979.

Keller, R. (2004). The computer aided resonance assignment tutorial, First edn (Goldau, Switzerland: Cantina Verlag).

Kohl, A., and Grutter, M.G. (2004). Fire and death: the pyrin domain joins the death-domain superfamily. Comptes rendus biologies *327*, 1077-1086.

Koshiba, T., Yasukawa, K., Yanagi, Y., and Kawabata, S. (2011). Mitochondrial membrane potential is required for MAVS-mediated antiviral signaling. Science signaling *4*, ra7.

Krauss, S., Zhang, C.Y., and Lowell, B.B. (2005). The mitochondrial uncoupling-protein homologues. Nature reviews Molecular cell biology *6*, 248-261.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature *227*, 680-685.

LeMaster, D.M. (1996). Structural determinants of the catalytic reactivity of the buried cysteine of Escherichia coli thioredoxin. Biochemistry *35*, 14876-14881.

Levitt, M.H. (2008). Spin Dynamics: Basics of Nuclear Magnetic Resonance., 2nd edn (Chichester: John Wiley & Sons, Ltd.).

Li, J., McQuade, T., Siemer, A.B., Napetschnig, J., Moriwaki, K., Hsiao, Y.S., Damko, E., Moquin, D., Walz, T., McDermott, A., *et al.* (2012). The RIP1/RIP3 necrosome forms a functional amyloid signaling complex required for programmed necrosis. Cell *150*, 339-350.

Li, X.D., Sun, L., Seth, R.B., Pineda, G., and Chen, Z.J. (2005). Hepatitis C virus protease NS3/4A cleaves mitochondrial antiviral signaling protein off the mitochondria to evade innate immunity. Proc Natl Acad Sci U S A *102*, 17717-17722.

Li, Y., Berthold, D.A., Frericks, H.L., Gennis, R.B., and Rienstra, C.M. (2007). Partial (13)C and (15)N chemical-shift assignments of the disulfide-bond-forming enzyme DsbB by 3D magic-angle spinning NMR spectroscopy. Chembiochem : a European journal of chemical biology *8*, 434-442.

Lin, S.C., Lo, Y.C., and Wu, H. (2010). Helical assembly in the MyD88-IRAK4-IRAK2 complex in TLR/IL-1R signalling. Nature *465*, 885-890.

Linge, J.P., Williams, M.A., Spronk, C.A., Bonvin, A.M., and Nilges, M. (2003). Refinement of protein structures in explicit solvent. Proteins *50*, 496-506.

Linser, R., Fink, U., and Reif, B. (2009). Probing surface accessibility of proteins using paramagnetic relaxation in solid-state NMR spectroscopy. Journal of the American Chemical Society *131*, 13703-13708.

Lipari, G., and Szabo, A. (1982a). Model-free approach to the interpretation of nuclear magnetic-resonance relaxation in macromolecules. 1. Theory and range of validity. J Am Chem Soc *104*, 4546-4559.

Lipari, G., and Szabo, A. (1982b). Model-free approach to the interpretation of nuclear magnetic-resonance relaxation in macromolecules. 2. Analysis of experimental results. J Am Chem Soc *104*, 4559-4570.

Liu, X.Y., Chen, W., Wei, B., Shan, Y.F., and Wang, C. (2011). IFN-induced TPR protein IFIT3 potentiates antiviral signaling by bridging MAVS and TBK1. J Immunol *187*, 2559-2568.

Loo, Y.M., Fornek, J., Crochet, N., Bajwa, G., Perwitasari, O., Martinez-Sobrido, L., Akira, S., Gill, M.A., Garcia-Sastre, A., Katze, M.G., *et al.* (2008). Distinct RIG-I and MDA5 signaling by RNA viruses in innate immunity. J Virol *82*, 335-345.

Loquet, A., Lv, G., Giller, K., Becker, S., and Lange, A. (2011). 13C spin dilution for simplified and complete solid-state NMR resonance assignment of insoluble biological assemblies. Journal of the American Chemical Society *133*, 4722-4725.

Loquet, A., Sgourakis, N.G., Gupta, R., Giller, K., Riedel, D., Goosmann, C., Griesinger, C., Kolbe, M., Baker, D., Becker, S., *et al.* (2012). Atomic model of the type III secretion system needle. Nature *486*, 276-279.

Lu, A., Magupalli, V.G., Ruan, J., Yin, Q., Atianand, M.K., Vos, M.R., Schroder, G.F., Fitzgerald, K.A., Wu, H., and Egelman, E.H. (2014). Unified polymerization mechanism for the assembly of ASC-dependent inflammasomes. Cell *156*, 1193-1206.

Ludtke, S.J., Baldwin, P.R., and Chiu, W. (1999). EMAN: semiautomated software for high-resolution single-particle reconstructions. Journal of structural biology *128*, 82-97.

Lundstrom, P., Teilum, K., Carstensen, T., Bezsonova, I., Wiesner, S., Hansen, D.F., Religa, T.L., Akke, M., and Kay, L.E. (2007). Fractional 13C enrichment of isolated carbons using [1-13C]- or [2-13C]-glucose facilitates the accurate measurement of dynamics at backbone Calpha and side-chain methyl positions in proteins. J Biomol NMR *38*, 199-212.

Mandel, A.M., Akke, M., and Palmer, A.G. (1995). Backbone dynamics of *Escherichia-coli* ribonuclease hi - correlations with structure and function in an active enzyme. J Mol Biol *246*, 144-163.

Markley, J.L., Ulrich, E.L., Westler, W.M., and Volkman, B.F. (2003). Macromolecular structure determination by NMR spectroscopy. Methods of biochemical analysis *44*, 89-113.

Martos, V., Castreno, P., Valero, J., and de Mendoza, J. (2008). Binding to protein surfaces by supramolecular multivalent scaffolds. Current opinion in chemical biology *12*, 698-706.

Masse, J.E., and Keller, R. (2005). AutoLink: Automated sequential resonance assignment of biopolymers from NMR data by relative-hypothesis-prioritization-based simulated logic. J Mag Res *174*, 133-151.

McDermott, A.E. (2004). Structural and dynamic studies of proteins by solid-state NMR spectroscopy: rapid movement forward. Curr Opin Struct Biol *14*, 554-561.

Meng, X., Zhao, G., and Zhang, P. (2011). Structure of HIV-1 capsid assemblies by cryo-electron microscopy and iterative helical real-space reconstruction. Journal of visualized experiments : JoVE.

Meylan, E., Curran, J., Hofmann, K., Moradpour, D., Binder, M., Bartenschlager, R., and Tschopp, J. (2005). Cardif is an adaptor protein in the RIG-I antiviral pathway and is targeted by hepatitis C virus. Nature *437*, 1167-1172.

Mibayashi, M., Martinez-Sobrido, L., Loo, Y.M., Cardenas, W.B., Gale, M., Jr., and Garcia-Sastre, A. (2007). Inhibition of retinoic acid-inducible gene I-mediated induction of beta interferon by the NS1 protein of influenza A virus. J Virol *81*, 514-524.

Michallet, M.C., Meylan, E., Ermolaeva, M.A., Vazquez, J., Rebsamen, M., Curran, J., Poeck, H., Bscheider, M., Hartmann, G., Konig, M., *et al.* (2008). TRADD protein is an essential component of the RIG-like helicase antiviral pathway. Immunity *28*, 651-661.

Morcombe, C.R., and Zilm, K.W. (2003). Chemical shift referencing in MAS solid state NMR. J Mag Res *162*, 479-486.

Moresco, E.M., Vine, D.L., and Beutler, B. (2011). Prion-like behavior of MAVS in RIG-I signaling. Cell Res *21*, 1643-1645.

Mukherjee, A., Morosky, S.A., Delorme-Axford, E., Dybdahl-Sissoko, N., Oberste, M.S., Wang, T., and Coyne, C.B. (2011). The coxsackievirus B 3C protease cleaves MAVS and TRIF to attenuate host type I interferon and apoptotic signaling. PLoS Pathog *7*, e1001311.

Mukherjee, M., Dutta, K., White, M.A., Cowburn, D., and Fox, R.O. (2006). NMR solution structure and backbone dynamics of domain III of the E protein of tick-borne Langat flavivirus suggests a potential site for molecular recognition. Protein Sci *15*, 1342-1355.

Mun, J.Y., Gulick, J., Robbins, J., Woodhead, J., Lehman, W., and Craig, R. (2011). Electron microscopy and 3D reconstruction of F-actin decorated with cardiac myosin-binding protein C (cMyBP-C). J Mol Biol *410*, 214-225.

Nagayama, K. (1981). Two-dimensional NMR spectroscopy: an application to the study of flexibility of protein molecules. Advances in biophysics *14*, 139-204.

Napetschnig, J., and Wu, H. (2013). Molecular basis of NF-kappaB signaling. Annual review of biophysics *42*, 443-468.

Okumura, A., Alce, T., Lubyova, B., Ezelle, H., Strebel, K., and Pitha, P.M. (2008). HIV-1 accessory proteins VPR and Vif modulate antiviral response by targeting IRF-3 for degradation. Virology *373*, 85-97.

Ou, H.D., Kwiatkowski, W., Deerinck, T.J., Noske, A., Blain, K.Y., Land, H.S., Soria, C., Powers, C.J., May, A.P., Shu, X., *et al.* (2012). A structural basis for the assembly and functions of a viral polymer that inactivates multiple tumor suppressors. Cell *151*, 304-319.

Palmer, A.G., 3rd, Kroenke, C.D., and Loria, J.P. (2001). Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules. Methods in enzymology *339*, 204-238.

Pan, K.M., Baldwin, M., Nguyen, J., Gasset, M., Serban, A., Groth, D., Mehlhorn, I., Huang, Z., Fletterick, R.J., Cohen, F.E., *et al.* (1993). Conversion of alpha-helices into beta-sheets features in the formation of the scrapie prion proteins. Proc Natl Acad Sci U S A *90*, 10962-10966.

Park, H.H., Lo, Y.C., Lin, S.C., Wang, L., Yang, J.K., and Wu, H. (2007a). The death domain superfamily in intracellular signaling of apoptosis and inflammation. Annu Rev Immunol *25*, 561-586.

Park, H.H., Logette, E., Raunser, S., Cuenin, S., Walz, T., Tschopp, J., and Wu, H. (2007b). Death domain assembly mechanism revealed by crystal structure of the oligomeric PIDDosome core complex. Cell *128*, 533-546.

Park, S., Juliana, C., Hong, S., Datta, P., Hwang, I., Fernandes-Alnemri, T., Yu, J.W., and Alnemri, E.S. (2013). The mitochondrial antiviral protein MAVS associates with NLRP3 and regulates its inflammasome activity. J Immunol *191*, 4358-4366.

Patel, D., Schultz, L.W., and Umland, T.C. (2013). Influenza A polymerase subunit PB2 possesses overlapping binding sites for polymerase subunit PB1 and human MAVS proteins. Virus research *172*, 75-80.

Paterson, Y., Englander, S.W., and Roder, H. (1990). An antibody binding site on cytochrome c defined by hydrogen exchange and two-dimensional NMR. Science *249*, 755-759.

Pauli, J., Baldus, M., van Rossum, B., de Groot, H., and Oschkinat, H. (2001). Backbone and side-chain 13C and 15N signal assignments of the alpha-spectrin SH3 domain by magic angle spinning solid-state NMR at 17.6 Tesla. Chembiochem : a European journal of chemical biology *2*, 272-281.

Peisley, A., Wu, B., Xu, H., Chen, Z.J., and Hur, S. (2014). Structural basis for ubiquitin-mediated antiviral signal activation by RIG-I. Nature *509*, 110-114.

Peisley, A., Wu, B., Yao, H., Walz, T., and Hur, S. (2013). RIG-I forms signaling-competent filaments in an ATP-dependent, ubiquitin-independent manner. Mol Cell *51*, 573-583.

Peng, J.W., and Wagner, G. (1992). Mapping of the spectral densities of N-H bond motions in eglin-C using heteronuclear relaxation experiments. Biochemistry *31*, 8571-8586.

Pines, A., Gibby, M.G., and Waugh, J.S. (1972). Proton-enhanced nuclear induction spectroscopy. A method for high resolution NMR of dilute spins in solids. J Chem Phys *56*, 1776-1777.

Pines, A., Gibby, M.G., and Waugh, J.S. (1973). Proton-enhanced NMR of dilute spins in solids. J Chem Phys *59*, 569-590.

Porcelli, A.M., Ghelli, A., Zanna, C., Pinton, P., Rizzuto, R., and Rugolo, M. (2005). pH difference across the outer mitochondrial membrane measured with a green fluorescent protein mutant. Biochemical and biophysical research communications *326*, 799-804.

Potter, J.A., Randall, R.E., and Taylor, G.L. (2008). Crystal structure of human IPS-1/MAVS/VISA/Cardif caspase activation recruitment domain. BMC Struct Biol *8*, 11.

Qiao, Q., Yang, C., Zheng, C., Fontan, L., David, L., Yu, X., Bracken, C., Rosen, M., Melnick, A., Egelman, E.H., *et al.* (2013). Structural architecture of the CARMA1/Bcl10/MALT1 signalosome: nucleation-induced filamentous assembly. Mol Cell *51*, 766-779.

Qin, H., Srinivasula, S.M., Wu, G., Fernandes-Alnemri, T., Alnemri, E.S., and Shi, Y. (1999). Structural basis of procaspase-9 recruitment by the apoptotic protease-activating factor 1. Nature *399*, 549-557.

Ramamoorthy, A., Wu, C.H., and Opella, S.J. (1999). Experimental aspects of multidimensional solid-state NMR correlation spectroscopy. J Mag Res *140*, 131-140.

Rand, U., Rinas, M., Schwerk, J., Nohren, G., Linnes, M., Kroger, A., Flossdorf, M., Kaly-Kullai, K., Hauser, H., Hofer, T., *et al.* (2012). Multi-layered stochasticity and paracrine signal propagation shape the type-I interferon response. Molecular systems biology *8*, 584.

Rebsamen, M., Vazquez, J., Tardivel, A., Guarda, G., Curran, J., and Tschopp, J. (2011). NLRX1/NOD5 deficiency does not affect MAVS signalling. Cell death and differentiation *18*, 1387.

Reed, J.C., Doctor, K.S., and Godzik, A. (2004). The domains of apoptosis: a genomics perspective. Science's STKE : signal transduction knowledge environment *2004*, re9.

Rees, I., Langley, E., Chiu, W., and Ludtke, S.J. (2013). EMEN2: an object oriented database and electronic lab notebook. Microscopy and microanalysis : the official journal of Microscopy Society of America, Microbeam Analysis Society, Microscopical Society of Canada *19*, 1-10.

Rieping, W., Habeck, M., Bardiaux, B., Bernard, A., Malliavin, T.E., and Nilges, M. (2007). ARIA2: automated NOE assignment and data integration in NMR structure calculation. Bioinformatics *23*, 381-382.

Ritter, C., Maddelein, M.L., Siemer, A.B., Luhrs, T., Ernst, M., Meier, B.H., Saupe, S.J., and Riek, R. (2005). Correlation of structural elements and infectivity of the HET-s prion. Nature *435*, 844-848.

Saha, S.K., Pietras, E.M., He, J.Q., Kang, J.R., Liu, S.Y., Oganesyan, G., Shahangian, A., Zarnegar, B., Shiba, T.L., Wang, Y., *et al.* (2006). Regulation of antiviral responses by a direct and specific interaction between TRAF3 and Cardif. The EMBO journal *25*, 3257-3263.

Saitô, H., Ando, I., and Naito, A. (2006). Solid State NMR Spectroscopy for Biopolymers : Principles and Applications. (Dordrecht: Springer).

Saito, T., and Gale, M., Jr. (2007). Principles of intracellular viral recognition. Curr Opin Immunol *19*, 17-23.

Schleucher, J., Schwendinger, M., Sattler, M., Schmidt, P., Schedletzky, O., Glaser, S.J., Sorensen, O.W., and Griesinger, C. (1994). A general enhancement scheme in heteronuclear multidimensional NMR employing pulsed-field gradients. Journal of biomolecular NMR *4*, 301-306.

Schmid, E.M., Ford, M.G., Burtey, A., Praefcke, G.J., Peak-Chew, S.Y., Mills, I.G., Benmerah, A., and McMahon, H.T. (2006). Role of the AP2 beta-appendage hub in recruiting partners for clathrin-coated vesicle assembly. PLoS biology *4*, e262.

Schwarzinger, S., Kroon, G.J., Foss, T.R., Chung, J., Wright, P.E., and Dyson, H.J. (2001). Sequence-dependent correction of random coil NMR chemical shifts. Journal of American Chemical Society *123*, 2970-2978.

Schwieters, C.D., Kuszewski, J.J., Tjandra, N., and Clore, G.M. (2003). The Xplor-NIH NMR molecular structure determination package. Journal of magnetic resonance *160*, 65-73.

Seth, R.B., Sun, L., and Chen, Z.J. (2006). Antiviral innate immunity pathways. Cell Res *16*, 141-147.

Seth, R.B., Sun, L., Ea, C.K., and Chen, Z.J. (2005). Identification and characterization of MAVS, a mitochondrial antiviral signaling protein that activates NF-kappaB and IRF 3. Cell *122*, 669-682.

Shahid, S.A., Bardiaux, B., Franks, W.T., Krabben, L., Habeck, M., van Rossum, B.J., and Linke, D. (2012a). Membrane-protein structure determination by solid-state NMR spectroscopy of microcrystals. Nature methods *9*, 1212-1217.

Shahid, S.A., Markovic, S., Linke, D., and van Rossum, B.J. (2012b). Assignment and secondary structure of the YadA membrane protein by solid-state MAS NMR. Scientific reports *2*, 803.

Shen, Y., Delaglio, F., Cornilescu, G., and Bax, A. (2009). TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. J Biomol NMR *44*, 213-223.

Shi, L., Kawamura, I., Jung, K.H., Brown, L.S., and Ladizhansky, V. (2011). Conformation of a seven-helical transmembrane photosensor in the lipid environment. Angewandte Chemie *50*, 1302-1305.

Shi, L., and Ladizhansky, V. (2012). Magic angle spinning solid-state NMR experiments for structural characterization of proteins. Methods Mol Biol *895*, 153-165.

Siegel, R.M., Martin, D.A., Zheng, L., Ng, S.Y., Bertin, J., Cohen, J., and Lenardo, M.J. (1998). Death-effector filaments: novel cytoplasmic structures that recruit caspases and trigger apoptosis. The Journal of cell biology *141*, 1243-1253.

Soares, F., Tattoli, I., Wortzman, M.E., Arnoult, D., Philpott, D.J., and Girardin, S.E. (2013). NLRX1 does not inhibit MAVS-dependent antiviral signalling. Innate immunity *19*, 438-448.

Song, T., Wei, C., Zheng, Z., Xu, Y., Cheng, X., Yuan, Y., Guan, K., Zhang, Y., Ma, Q., Shi, W., *et al.* (2010). c-Abl tyrosine kinase interacts with MAVS and regulates innate immune response. FEBS letters *584*, 33-38.

Spera, S., and Bax, A. (1991). Empirical correlation between protein backbone conformation and C-alpha and C-beta C-13 nuclear-magnetic-resonance chemical-shifts. J Am Chem Soc *113*, 5490-5492.

Stevens, T.J., Fogh, R.H., Boucher, W., Higman, V.A., Eisenmenger, F., Bardiaux, B., van Rossum, B.J., Oschkinat, H., and Laue, E.D. (2011). A software framework for analysing solid-state MAS NMR data. J Biomol NMR *51*, 437-447.

Strober, W., Murray, P.J., Kitani, A., and Watanabe, T. (2006). Signalling pathways and molecular interactions of NOD1 and NOD2. Nat Rev Immunol *6*, 9-20.

Subramanian, N., Natarajan, K., Clatworthy, M.R., Wang, Z., and Germain, R.N. (2013). The adaptor MAVS promotes NLRP3 mitochondrial localization and inflammasome activation. Cell *153*, 348-361.

Suen, D.F., Norris, K.L., and Youle, R.J. (2008). Mitochondrial dynamics and apoptosis. Genes & development *22*, 1577-1590.

Sun, Q., Sun, L., Liu, H.H., Chen, X., Seth, R.B., Forman, J., and Chen, Z.J. (2006). The specific and essential role of MAVS in antiviral innate immune responses. Immunity *24*, 633-642.

Takamatsu, S., Onoguchi, K., Onomoto, K., Narita, R., Takahasi, K., Ishidate, F., Fujiwara, T.K., Yoneyama, M., Kato, H., and Fujita, T. (2013). Functional characterization of domains of IPS-1 using an inducible oligomerization system. PloS one *8*, e53578.

Takegoshi, K., Nakamura, S., and Terao, T. (2001). C^{13}-H^{1} dipolar-assisted rotational resonance in magic-angle spinning NMR. Chem Phys Lett *344*, 631-637.

Takeuchi, O., and Akira, S. (2009). Innate immunity to virus infection. Immunol Rev *227*, 75-86.

Takeuchi, O., and Akira, S. (2010). Pattern recognition receptors and inflammation. Cell *140*, 805-820.

Tang, E.D., and Wang, C.Y. (2009). MAVS self-association mediates antiviral innate immune signaling. J Virol *83*, 3420-3428.

Thompson, L.K. (2002). Solid-state NMR studies of the structure and mechanisms of proteins. Curr Opin Struct Biol *12*, 661-669.

Tinel, A., and Tschopp, J. (2004). The PIDDosome, a protein complex implicated in activation of caspase-2 in response to genotoxic stress. Science *304*, 843-846.

Torizawa, T., Shimizu, M., Taoka, M., Miyano, H., and Kainosho, M. (2004). Efficient production of isotopically labeled proteins by cell-free synthesis: a practical protocol. Journal of biomolecular NMR *30*, 311-325.

Uematsu, S., and Akira, S. (2007). Toll-like receptors and Type I interferons. J Biol Chem *282*, 15319-15323.

Ulmer, T.S., Campbell, I.D., and Boyd, J. (2002). The effects of dissolved oxygen upon amide proton relaxation and chemical shift in a perdeuterated protein. Journal of magnetic resonance *157*, 181-189.

van Eck, E.R., and Veeman, W.S. (1993). Spin density description of rotational-echo double-resonance, transferred-echo double-resonance and two-dimensional transferred-echo double-resonance solid state nuclear magnetic resonance. Solid state nuclear magnetic resonance *2*, 307-315.

Varga, Z.T., Grant, A., Manicassamy, B., and Palese, P. (2012). Influenza virus protein PB1-F2 inhibits the induction of type I interferon by binding to MAVS and decreasing mitochondrial membrane potential. J Virol *86*, 8359-8366.

Vaughn, D.E., Rodriguez, J., Lazebnik, Y., and Joshua-Tor, L. (1999). Crystal structure of Apaf-1 caspase recruitment domain: an alpha-helical Greek key fold for apoptotic signaling. J Mol Biol *293*, 439-447.

Verardi, R., Traaseth, N.J., Masterson, L.R., Vostrikov, V.V., and Veglia, G. (2012). Isotope labeling for solution and solid-state NMR spectroscopy of membrane proteins. Advances in experimental medicine and biology *992*, 35-62.

Vitour, D., Dabo, S., Ahmadi Pour, M., Vilasco, M., Vidalain, P.O., Jacob, Y., Mezel-Lemoine, M., Paz, S., Arguello, M., Lin, R., *et al.* (2009). Polo-like kinase 1 (PLK1) regulates interferon (IFN) induction by MAVS. J Biol Chem *284*, 21797-21809.

Voortman, L.M., Franken, E.M., van Vliet, L.J., and Rieger, B. (2012). Fast, spatially varying CTF correction in TEM. Ultramicroscopy *118*, 26-34.

Vranken, W.F., Boucher, W., Stevens, T.J., Fogh, R.H., Pajon, A., Llinas, M., Ulrich, E.L., Markley, J.L., Ionides, J., and Laue, E.D. (2005). The CCPN data model for NMR spectroscopy: development of a software pipeline. Proteins *59*, 687-696.

Vriend, G. (1990). WHAT IF: a molecular modeling and drug design program. Journal of molecular graphics *8*, 52-56, 29.

Wang, L., Yang, J.K., Kabaleeswaran, V., Rice, A.J., Cruz, A.C., Park, A.Y., Yin, Q., Damko, E., Jang, S.B., Raunser, S., *et al.* (2010). The Fas-FADD death domain complex structure reveals the basis of DISC assembly and disease mutations. Nature structural & molecular biology *17*, 1324-1329.

Wang, Y.J., and Jardetzky, O. (2002a). Probability-based protein secondary structure identification using combined NMR chemical-shift data. Protein Sci *11*, 852-861.

Wang, Y.J., and Jardetzky, O. (2002b). Probability-based protein secondary structure identification using combined NMR chemical-shift data. Protein Science *11*, 852-861.

Wasmer, C., Lange, A., Van Melckebeke, H., Siemer, A.B., Riek, R., and Meier, B.H. (2008). Amyloid fibrils of the HET-s(218-289) prion form a beta solenoid with a triangular hydrophobic core. Science *319*, 1523-1526.

Wei, C., Ni, C., Song, T., Liu, Y., Yang, X., Zheng, Z., Jia, Y., Yuan, Y., Guan, K., Xu, Y., *et al.* (2010). The hepatitis B virus X protein disrupts innate immunity by downregulating mitochondrial antiviral signaling protein. J Immunol *185*, 1158-1168.

Wen, C., Yan, Z., Yang, X., Guan, K., Xu, C., Song, T., Zheng, Z., Wang, W., Wang, Y., Zhao, M., *et al.* (2012). Identification of tyrosine-9 of MAVS as critical target for inducible phosphorylation that determines activation. PloS one *7*, e41687.

West, A.P., Shadel, G.S., and Ghosh, S. (2011). Mitochondria in innate immune responses. Nat Rev Immunol *11*, 389-402.

Wider, G. (2000). Structure determination of biological macromolecules in solution using nuclear magnetic resonance spectroscopy. BioTechniques *29*, 1278-1282, 1284-1290, 1292 passim.

Wider, G., and Wüthrich, K. (1999). NMR spectroscopy of large molecules and multimolecular assemblies in solution. Curr Opin Struct Biol *9*, 594-601.

Wishart, D.S., and Nip, A.M. (1998). Protein chemical shift analysis: a practical guide. Biochemistry and cell biology = Biochimie et biologie cellulaire *76*, 153-163.

Wishart, D.S., and Sykes, B.D. (1994). The ^{13}C chemical-shift index: - A simple method for the identification of protein secondary structure using ^{13}C chemical-shift data. Journal of biomolecular NMR *4*, 171-180.

Wright, E.R. (2010). Correlative structural biology: how to investigate the fine details of viral structure. Viruses *2*, 107-110.

Wu, B., Peisley, A., Richards, C., Yao, H., Zeng, X., Lin, C., Chu, F., Walz, T., and Hur, S. (2013). Structural basis for dsRNA recognition, filament formation, and antiviral signal activation by MDA5. Cell *152*, 276-289.

Wu, H. (2013). Higher-order assemblies in a new paradigm of signal transduction. Cell *153*, 287-292.

Wüthrich, K. (1986). NMR of Proteins and Nucleic Acids (New York: Joh Wiley & Sons).

Xiao, T., Towb, P., Wasserman, S.A., and Sprang, S.R. (1999). Three-dimensional structure of a complex between the death domains of Pelle and Tube. Cell *99*, 545-555.

Xu, H., He, X., Zheng, H., Huang, L.J., Hou, F., Yu, Z., de la Cruz, M.J., Borkowski, B., Zhang, X., Chen, Z.J., *et al.* (2014). Structural basis for the prion-like MAVS filaments in antiviral innate immunity. eLife *3*, e01489.

Xu, L., Xiao, N., Liu, F., Ren, H., and Gu, J. (2009). Inhibition of RIG-I and MDA5-dependent antiviral response by gC1qR at mitochondria. Proc Natl Acad Sci U S A *106*, 1530-1535.

Xu, L.G., Wang, Y.Y., Han, K.J., Li, L.Y., Zhai, Z., and Shu, H.B. (2005). VISA is an adapter protein required for virus-triggered IFN-beta signaling. Mol Cell *19*, 727-740.

Yang, K., Shi, H., Qi, R., Sun, S., Tang, Y., Zhang, B., and Wang, C. (2006). Hsp90 regulates activation of interferon regulatory factor 3 and TBK-1 stabilization in Sendai virus-infected cells. Molecular biology of the cell *17*, 1461-1471.

Yang, Y., Liang, Y., Qu, L., Chen, Z., Yi, M., Li, K., and Lemon, S.M. (2007). Disruption of innate immunity due to mitochondrial targeting of a picornaviral protease precursor. Proc Natl Acad Sci U S A *104*, 7253-7258.

Yasukawa, K., Oshiumi, H., Takeda, M., Ishihara, N., Yanagi, Y., Seya, T., Kawabata, S., and Koshiba, T. (2009). Mitofusin 2 inhibits mitochondrial antiviral signaling. Science signaling *2*, ra47.

Yin, Q., Lin, S.C., Lamothe, B., Lu, M., Lo, Y.C., Hura, G., Zheng, L., Rich, R.L., Campos, A.D., Myszka, D.G., *et al.* (2009). E2 interaction and dimerization in the crystal structure of TRAF6. Nature structural & molecular biology *16*, 658-666.

Yoneyama, M., and Fujita, T. (2008). Structural mechanism of RNA recognition by the RIG-I-like receptors. Immunity *29*, 178-181.

Yoneyama, M., and Fujita, T. (2009). RNA recognition and signal transduction by RIG-I-like receptors. Immunol Rev *227*, 54-65.

Yoneyama, M., Kikuchi, M., Matsumoto, K., Imaizumi, T., Miyagishi, M., Taira, K., Foy, E., Loo, Y.M., Gale, M., Jr., Akira, S., *et al.* (2005). Shared and unique functions of the DExD/H-box helicases RIG-I, MDA5, and LGP2 in antiviral innate immunity. J Immunol *175*, 2851-2858.

You, F., Sun, H., Zhou, X., Sun, W., Liang, S., Zhai, Z., and Jiang, Z. (2009). PCBP2 mediates degradation of the adaptor MAVS via the HECT ubiquitin ligase AIP4. Nat Immunol *10*, 1300-1308.

Young, J.C., Hoogenraad, N.J., and Hartl, F.U. (2003). Molecular chaperones Hsp90 and Hsp70 deliver preproteins to the mitochondrial import receptor Tom70. Cell *112*, 41-50.

Zeng, W., Sun, L., Jiang, X., Chen, X., Hou, F., Adhikari, A., Xu, M., and Chen, Z.J. (2010). Reconstitution of the RIG-I pathway reveals a signaling role of unanchored polyubiquitin chains in innate immunity. Cell *141*, 315-330.

Zhang, P., Meng, X., and Zhao, G. (2013). Tubular crystals and helical arrays: structural determination of HIV-1 capsid assemblies using iterative helical real-space reconstruction. Methods Mol Biol *955*, 381-399.

Zhao, Y., Sun, X., Nie, X., Sun, L., Tang, T.S., Chen, D., and Sun, Q. (2012). COX5B regulates MAVS-mediated antiviral signaling through interaction with ATG5 and repressing ROS production. PLoS Pathog *8*, e1003086.

Zheng, C., Kabaleeswaran, V., Wang, Y., Cheng, G., and Wu, H. (2010). Crystal structures of the TRAF2: cIAP2 and the TRAF1: TRAF2: cIAP2 complexes: affinity, specificity, and regulation. Mol Cell *38*, 101-113.

Zhong, B., Yang, Y., Li, S., Wang, Y.Y., Li, Y., Diao, F., Lei, C., He, X., Zhang, L., Tien, P., *et al.* (2008). The adaptor protein MITA links virus-sensing receptors to IRF3 transcription factor activation. Immunity *29*, 538-550.

Zhou, S., Cerny, A.M., Zacharia, A., Fitzgerald, K.A., Kurt-Jones, E.A., and Finberg, R.W. (2010). Induction and inhibition of type I interferon responses by distinct components of lymphocytic choriomeningitis virus. J Virol *84*, 9452-9462.

Zou, J., Chang, M., Nie, P., and Secombes, C.J. (2009). Origin and evolution of the RIG-I like RNA helicase gene family. BMC Evol Biol *9*, 85.